Information
Resource
Center

MD6
8.2

CLEARINGHOUSE

II

AN URBAN INSTITUTE PAPER ON HEALTH

Controlling Medicaid Utilization Patterns

JOHN HOLAHAN
BRUCE STUART

RA
412
.4
H64
1977

BOARD OF TRUSTEES

William D. Ruckelshaus,
  Chairman
Warren E. Buffett
Anthony Downs
John H. Filer
Eugene G. Fubini
William Gorham
Katharine Graham
Vernon E. Jordan, Jr.
Richard Llewelyn-Davies
Bayless A. Manning
Robert S. McNamara
Arjay Miller
Lois D. Rice
Herbert E. Scarf
Henry B. Schacht
Irving S. Shapiro
Franklin A. Thomas
John G. Veneman
James Vorenberg

INSTITUTE OFFICERS

William Gorham, *President*
Robert Harris, *Senior Vice-President*
Harold W. Guthrie, *Vice-President for Research*
Herman T. Boland, Jr., *Vice-President for Administration*
Edward E. Wallace, *Controller*

The Urban Institute is a nonprofit research corporation established in 1968 to study problems of the nation's urban communities. Independent and nonpartisan, the Institute responds to current needs for disinterested analyses and basic information and attempts to facilitate the application of this knowledge. As part of this effort, it cooperates with federal agencies, states, cities, associations of public officials, the academic community and other sectors of the general public.

The Institute's research findings and a broad range of interpretive viewpoints are published as an educational service. Conclusions expressed in Institute publications are those of the authors and do not necessarily reflect the views of other staff members, officers or trustees of the Institute, or of organizations which provide funds toward support of Institute studies.

These research findings are made available in three series of publications: Books and Reports, Papers, and Reprints. A current publications list is available on request.

Publications Office

The Urban Institute
2100 M Street, N.W.
Washington, D.C. 20037

RA
412.4
.H64
1977

CLEARINGHOUSE

Controlling Medicaid Utilization Patterns

JOHN HOLAHAN
BRUCE STUART

986-15

June 1977

THE URBAN INSTITUTE
WASHINGTON, D.C.

The research forming the basis for this report was funded in part by the National Center for Health Services Research of the U.S. Department of Health, Education and Welfare through Contract HRA-230-75-0122. Additional funds were provided by the Ford Foundation.

The interpretations or conclusions are those of the authors and should not be attributed to The Urban Institute, its trustees, or to other organizations that support its research.

ISBN 87766-196-0

UI 986-15

PLEASE REFER TO URI 17900 WHEN ORDERING

Available from:

Publications Office
The Urban Institute
2100 M Street, N.W.
Washington, D.C. 20037

List price: $4.00

A/77/1M

TABLE OF CONTENTS

Acknowledgments v

Preface vii

CHAPTER I. COST SHARING 1
John Holahan

A. Introduction. 1

B. Physician Services. 5

C. Hospitals 16

D. Prescription Drugs. 19

E. Conclusions 24

CHAPTER II. UTILIZATION CONTROLS 29
Bruce Stuart

A. Introduction. 29

B. Requirements for Effective Utilization Control. 30

C. The Extent and Cost of Unnecessary and Inappropriate Utilization 34

1. Medically Unnecessary Services. 35
2. Medically Inappropriate Utilization 38
3. Socially Inappropriate Utilization. 42
4. Fraud and Abuse 44
5. Measuring the Cost of Misutilization. 45
6. Causal Factors in Misutilization. 49

D. Classifying Utilization Controls. 52

E. Predelivery Controls. 53

1. Prior Review Mechanisms 53
2. Prior Authorization 65

F. Concurrent Review 81

G. Post-Treatment Review 94

1. Institutional Utilization Review. 97
2. PSRO Profile Analysis 99
3. Surveillance and Utilization Review Systems 102

H. Conclusions 108

Notes

Notes to Chapter I. 115
Notes to Chapter II 117

ACKNOWLEDGMENTS

The authors wish to thank the following individuals for their helpful criticisms and comments: Bruce Spitz, Jack Hadley, David Salkever and Terry McCrary.

PREFACE

This paper is the second of a four-part series on alternative strategies for controlling the costs of state Medicaid programs. The preparation of these papers was funded by the National Center for Health Services Research, as part of a larger effort to analyze the effect that the 1974-1975 recession had on the states' ability to finance health services for the poor. The Medicaid Cost Containment Series is a systematic examination of (1) state options which are permitted by federal regulations and (2) state options which would improve the efficiency and effectiveness of the program. These alternatives do not necessarily overlap. Indeed, federal regulations often preclude rational state behavior. We therefore set forth cost containment issues and options regarding Medicaid eligibility, benefits, provider reimbursement and utilization controls. We then address possible changes in federal regulations and financing mechanisms which might improve state management of Medicaid.

The first chapter, Cost Sharing, examines the arguments for and against cost sharing as a cost control device available to the states. States are currently permitted the use of cost sharing, the payment by the recipient of some portion of the bill for a service at time of use, as a cost control device. However, use must be limited to nominal copayments on optional services for cash assistance recipients and on all services for the medically needy. The most common rationale for the use of cost sharing is that it will introduce cost consciousness on the part of recipients, thereby discouraging the use of unnecessary services, the excessive prescribing of drugs, etc.

The chapter suggests that cost sharing, if imposed mostly on routine services, can be a fairly equitable approach to reducing program costs, when compared to other forms of control (e.g., elimination of services, limits on hospital days) where the burden falls on those with severe but specific needs for these services. Equity can be increased if certain essential services are completely excluded from copayment or if cash assistance grants are increased at the same time.

The evidence on the efficiency aspect of cost sharing is less clear. The limited empirical research on the effects of copayment on the poor suggests that there may be a limiting effect on utilization of physicians' services, at least in the short run (which may well be important to a state faced with severe budget problems). In the long run, however, provider response to a decrease in patient-initiated demands for care may be increased demand creation. Whether copayment on physicians' services has adverse effects on patient health, or whether the deterrent effect eventually wears off, has not been substantiated. The paper argues against the use of cost sharing on hospital inpatient care. Cost sharing on hospital stays could become a serious burden on patients. Moreover, hospitals can too easily respond by leaving the charge uncollected and manipulating the charge structure on lengths of stay to recoup lost funds. Copayment for prescription drugs would probably have some deterrent effect on the use of unnecessary drugs. The actual results, however, may be reduced pharmacy participation and increase in size of prescriptions. It may also reduce use of necessary prescriptive drugs, resulting in use of other, more expensive services.

Finally, the paper suggests that cost sharing, if carefully applied, may offer a powerful technique for policing the system and helping to limit abuses, such as unnecessary diagnostic services or unnecessary referrals. Medicaid patients presently have no financial incentive to oppose such practices; the entire burden of limiting the abusive practices of physicians rests with the administrative machinery of the state.

The second chapter, Utilization Controls, provides a comprehensive examination of utilization control strategies available to the states. Utilization control is defined as any regulatory mechanism which materially affects the behavior of providers and/or patients in terms of the quantity, quality or type of medical service utilized. The need for utilization control under Medicaid is based on two assumptions: first, that there is a level of serious misutilization in current systems of medical care delivery, and second, that regulatory rather than market reform is the appropriate public response. The paper examines at length both the limitations on and the potential for cost containment through utilization control. The requirements for effective utilization control are delineated as (a) appropriate targeting, (b) enforceability and (c) efficient administration.

The chapter examines various methods of utilization controls, grouping them on the basis of the timing of the control relative to service delivery; in other words, on whether the control is applied before, during or after treatment. Predelivery controls include prior review mechanisms, including either voluntary or mandatory medical review before services will be considered for payment, and prior authorization, which requires a formal authorization in addition to the review procedure. Concurrent review is only feasible for extended treatment requirements and, for all practical purposes, has been limited to inpatient hospital services. Post-treatment review differs fundamentally from the other types in that it comes into play only after treatment is rendered, and consists of institutional utilization review, PSRO profile analysis, and surveillance and utilization review (S/UR) systems.

Definition and objectives, current regulatory status, evidence of effectiveness and adoption implications are all discussed in some detail for each of the review mechanisms. Of the mechanisms, prior authorization, where appropriately applied, has certain very attractive features. It can be targeted at high-volume providers or expensive services or both, and it is relatively inexpensive to administer. Surveillance and utilization review systems can also be critical weapons for states in their cost control efforts, as they help to establish uniform standards of care, screen exceptions, target direct intervention mechanisms when standards are not met and provide a deterrent to future misutilization and abuse.

CHAPTER I

COST SHARING

A. Introduction

Cost sharing, the payment of some portion of the bill for a service at time of use, has been a widely discussed and highly controversial issue in Medicaid. Prior to the 1972 Amendments to the Social Security Act, PL92-603, states could not use cost sharing as a cost control device. That legislation permitted states to impose nominal copayments on optional services for cash assistance recipients and on all services for the medically needy. States are not permitted to impose any form of cost sharing on cash assistance recipients for such services as physician visits or laboratory services. States interested in imposing cost sharing on those services are limited to applying it only to the non-institutionalized medically needy, a relatively small group, and must require providers to make a distinction between copayers and non-copayers. Other administrative constraints make copayments directly proportionate to the state's payment of a service and set a series of screens for the application of coinsurance and deductibles.

As a result, no states currently use cost sharing for physician or hospital outpatient services. Ten states currently have cost sharing for prescription drugs. Several states responded to the recent recession by adopting copayments for optional services. For example, Georgia imposed copayments on ambulance services, prescription refills, durable medical equipment and prosthetic devices, and attempted, through a special HEW waiver, a $25.00 deductible on hospital stays and copayments for physician services and outpatient services. New Jersey introduced a twenty-five cent prescription charge, and Alabama required a fifty cent prescription drug refill copayment. Virginia imposed a two dollar charge for new glasses or repair. Other states contemplated similar actions.

Discussions of proposed cost sharing in Medicaid most frequently focus on the use of copayments for various services. However, there are three relevant types of cost sharing: copayments, coinsurance and deductibles. With a deductible, the patient would pay all costs up to a certain level, such as the first twenty-five dollars of a hospital stay. In a coinsurance arrangement, the patient's liability is fixed at some percentage of the bill. Under a copayment scheme, which is the most commonly used by state Medicaid programs, a fixed charge per unit of service is assessed, e.g., two dollars per office visit.

The most common rationale for use of cost sharing is that it will introduce cost consciousness on the part of recipients, thereby discouraging use of unnecessary services while encouraging search for more efficient low-cost producers. When Medicaid pays all of a physician's or hospital's bill, the services are free to the recipient; recipients can be expected to use more services than if a price is attached. Their demand for free medical services will only be limited by transportation costs, the cost of lost time (e.g., wages, leisure) and the pain sometimes associated with the consumption of medical care. The use of some form of cost sharing will reduce the amount of care demanded, unless recipients are totally unresponsive to price (e.g., demand is perfectly inelastic with respect to price).

It is also claimed that cost sharing may make the system easier to police, offering a powerful technique for limiting such abuses as the provision of unnecessary laboratory tests and X-rays, referrals to other practitioners within a clinic and excessive prescribing of drugs. Such abuses received nationwide publicity with the Moss Committee hearings on fraud and abuse among practitioners in Medicaid. That such abuses are easily perpetuated when a patient bears no share of the cost for services has been largely overlooked. The absence of cost sharing places an enormous administrative responsibility on the state, one which states have proven unable to bear at present.

Third, it is argued that cost sharing would make it feasible to finance a broader benefit package, including items, such as dental care, which are postponable for many but which impose severe hardships on others. In the extreme case that demand is totally unresponsive to charges, the only effect of cost sharing would be to distribute some of the burden of financing the program to the recipients. States could, in principle, use the recipient payments to add to the benefit package.

Finally, it can be argued that, in comparison to the most prevalent form of utilization control exercised by the states during the recent recession (i.e., elimination of optional services and limits on hospital days), cost sharing is more equitable in that the burden is somewhat broadly shared rather than imposed on those with severe but specific needs for services. Individuals need not bear severe financial hardship if copayment provisions are designed to apply only to most of the routine services while the program continues to protect the individual against large expenses for a broad range of health services.

While cost sharing can, potentially, offer a fairly equitable approach to reducing program costs, there is widespread disagreement on its efficiency. Many argue that deductibles and coinsurance perform the allocation function in an inefficient manner. That is, patients have insufficient knowledge to make rational calculations of the benefits and costs of their choices, while physicians and other providers who presumably do possess adequate information are only indirectly affected, if at all, by the prices facing consumers. Some have argued that physicians respond to reductions in patient-initiated demands by providing more diagnostic services, more follow-up visits, etc. It has also been argued that copayments in Medicaid will meet with another type of reaction--substantial resistance from providers and reduced participation in the program, thereby limiting the access of patients to needed care. Another result may be that cost sharing will deter early utilization of necessary services, thus adversely affecting health and leading to higher utilization rates in the long run. These critics urge that other forms of cost control be imposed, particularly those which impinge on the physician. Finally, it is also argued that, if cost sharing is employed to reduce the impact of the Medicaid program on federal and state budgets, the objective can be equally well achieved in other ways, i.e., through utilization review or reimbursement controls which would avoid the necessity of imposing any financial burdens on the sick during a time of illness.

A fundamental problem in assessing the effectiveness of copayments is the fact that relatively little information exists concerning the level and pattern of utilization brought about by some form of cost sharing.

Loss of income from work, travel costs, fear, ignorance and availability of facilities are all factors associated with seeking medical attention. Research that isolates the magnitude and force of these other considerations from the influence of copayments on utilization is limited. In the following pages, we will summarize available evidence on the impact of cost sharing on physician services, hospitalization and prescription drugs.

B. Physician Services

No states currently impose cost sharing on physician services. Nonetheless, there has been and continues to be considerable interest in it as a cost control mechanism. Some limited evidence concerning copayment as a mechanism for containing the use of physician services is available. The Canadian province of Saskatchewan provided one of the first opportunities to examine the overall impact of copayment on the utilization of certain specific medical care services.[1] In 1946, a prepaid medical service plan was established in one part of the province; it provided comprehensive care to 52,000 persons residing in a 12,000 square mile area. The residents were also covered by a province-wide hospitalization scheme of unlimited ward care, financed by the provincial government. The doctors were paid on a fee-for-service basis at a stipulated percentage of the fee schedule of the provincial College of Physicians and Surgeons. The thirteen hospitals in the area provided an ample supply of beds to handle all acute care patients in need of admission.

In 1953, seven years after the plan's inception, physicians and administrators agreed to levy a copayment on what were believed to be excessively demanded home and office visits. Doctors were to collect a one dollar fee from their patients for all office calls, minor surgery in the office and first visits for maternity care, a two dollar charge for at-home day calls and a three dollar copayment for at-home night calls. Because hospital visits were perceived as physician-induced, not patient-demanded, no copayment was applied on these services.

The impact of the copayment scheme was immediate and substantial. By the end of 1954, office visits had decreased by almost 20 percent. Home visits, perhaps reflecting the higher copayment, experienced an even sharper decline, falling by approximately 50 percent. The overall program effect, however, was mitigated by an increase in hospital visits, which rose by about 15 percent. The most significant increase occurred in the incidence of minor surgery, performed in hospitals, which jumped from 168 per thousand persons in 1952 to 199 per thousand persons in 1954.

Moreover, with the passage of time there were indications that physicians did not always attempt to collect the copayment charges; to the extent that they did, it was not determined how widespread nonpayment had become. This phenomenon coincided with a reversal of the downward trend in home and office visits that began in 1955. By the end of 1959, office calls almost matched the figures prevailing prior to copayment. Home visits, also, showed a significant increase from the low incidence attained in 1953 and 1954. These services stood in sharp contrast to hospital admissions which sustained throughout the decade the high levels first experienced following the adoption of copayment.

It is clear in the light of later evidence that the copayment plan gradually lost its deterrent effectiveness. What is still in doubt, however, are the precise reasons for the overall lackluster impact of copayment on utilization. The author of a report on the plan speculates that it may be that "to replace the lost income from a larger amount of uncollected coinsurance charges, the doctors were financially encouraged to supply more follow-up visits and services than before."[2] Unfortunately, no data were gathered to measure the change (if any) in the ratio of repeat visits to first visits in an illness. Other possible explanations include the effect of inflation on the real value of fixed copayments and the effect of increases in incomes over time.

Some years later, Saskatchewan provided another opportunity for analysis of the copayment effect on utilization. In 1962, the province developed a medical care insurance plan which provided comprehensive coverage to virtually all of its residents. The government program incorporated few exclusions and paid physicians on a fee-for-service basis.

In 1968, the provincial government instituted a "utilization" (i.e., a copayment) fee which permitted the physician to charge $1.50 for office visits and $2.00 for home, emergency or hospital outpatient visits. Beck[3] segregated the affected population according to economic status, i.e., the "poor" and the "non-poor." The results indicated that the copayment effected, on the average, a decline in services to poor families of approximately 18 percent in 1968. When the analysis included age of head and family size, in addition to income, the reduction in utilization of services was about 24 percent for aged and large families.

This contrasted dramatically with the rest of the population, which experienced a medical service reduction estimated at 6 to 7 percent.

The study found clear evidence of a pure price effect on consumers, precipitated by the introduction of a copayment and resulting in a decline in the quantity of services taken. Evidence from the Saskatchewan Insurance Commission of a 8.9 percent increase between 1969 and 1970 in adjusted per capita utilization of services, however, supports the earlier contention that the effect of copayments is transitory.[4] In his study, Beck raises the possibility of a positive supply response by physicians to copayment; increased costs of collection and bad debts are absorbed by the physician and result in a greater quantity of medical services supplied. This effect is possible if physicians have income targets and exercise their influence over demand to meet that objective. He provided evidence that physician-initiated services (i.e., laboratory services, major surgery) declined less than did patient-initiated services, as well as evidence of a substitution of complete examinations for less expensive regional examinations.

No firm conclusions can be drawn for the long-run effects of copayments from the Saskatchewan experience. First, the effects of inflation on the real value of fixed-dollar copayments and the effect of increases in incomes were again present. Second, it should be emphasized that a copayment charge was permitted, not required, by the doctors. No attempt was made to determine the extent to which a copayment was in fact extracted from the patients.

Scitovsky and Snyder,[5] followed by Phelps and Newhouse,[6] analyzed the effects of coinsurance on the demand for physician services by the employees and faculty of Stanford University. In 1966, no costs were charged for clinic services to members of the Palo Alto Group Health Plan. In the following year, however, to offset considerable losses, a 25 percent coinsurance rate was established for outpatient medical, surgical and ancillary services. The data showed that physician visits and expenditures declined by approximately 25 percent; utilization of ancillary services also declined by one-half. Home visits decreased the most (51 percent); office visits about average (25 percent); and hospital non-elective surgery the least (4.8 percent).

The Phelps and Newhouse study also examined the effects of time prices on demand for medical care, i.e., the extent to which time played a role in influencing the demand of various groups of subscribers. The time prices of medical care services are lower for low-income individuals, particularly female dependents. Thus, a given increase in a money price will have a large effect for those with low time prices. Phelps and Newhouse found, as expected, that utilization among different groups varied inversely with the price of time they were assumed to face. The decline in visits brought about by coinsurance was greater for female dependents than for other categories.

A recent follow-up study by Scitovsky and McCall (formerly Snyder) contradicts the Canadian evidence on the long-run effects of coinsurance. The study examined physician service utilization rates for 1972 for the original

1966-1968 cohort and found no evidence of an upward trend in physician utilization rates. There are three important differences between the Palo Alto and the Canadian evidence. First, copayments were used in Canada, while coinsurance was applied in Palo Alto. Thus, coinsurance payments rise as physician fees rise, reducing the effect of inflation on the real value of patient cost sharing. Second, the coinsurance in the Palo Alto Group Health Plan was collected from patients, whereas there is some question about the level of effort to collect copayments in Saskatchewan. Third, physicians in the Palo Alto Group Health Plan are on salary; they would not benefit from demand creation as would the physicians in Saskatchewan, who were reimbursed on a fee-for-service basis.

Perhaps the most ambitious experiment to date on the impact of copayment on utilization was conducted by the State of California with its Medi-Cal (Medicaid) program. Between January 1, 1972, and June 30, 1973, approximately 30 percent of those eligible for Medi-Cal were required to pay a dollar for each of their first two visits to providers each month and fifty cents for the first two prescriptions they had filled each month. The copayment was made by the Medi-Cal recipients directly to the provider and pharmacist upon delivery of service. The charge requirement was applicable only to those who had property and income above a defined level.

The data gathered from this experiment were consistent to a large extent with the findings established in previous studies conducted in other parts of the country and among different socio-economic groups. The Medi-Cal experiment showed that copayment tended to reduce overall beneficiary utilization; this reduction, however, took place primarily in medical services for which treatment was viewed as not serious or, at least, postponable. Thus, the study revealed that a dramatic reduction took place in the utilization of preventive services. Within the AFDC category, the immunization rates dropped by 45 percent, "total obstetrical care" (which included prenatal care) was reduced by 58 percent, and pap smears (for detection of cervical cancer) were lowered by 22 percent.

Finally, a modest reduction was found in the use of physician services. Copayer utilization was about 17 percent lower than non-copayer in the AFDC category, 1 percent higher among the aged, and 4 percent higher in the blind and disabled categories. All aid categories combined, however, visited their physicians less often and received fewer services during each visit. The lower rate of services received ranged from over 30 percent lower in AFDC, 11 percent lower for the aged, and 2 percent lower for the blind and disabled.

It should be stressed that, though the experiment indicated a high responsiveness in utilization due to copayment, the California study was flawed by several serious methodological and structural shortcomings. First, in the three months before the implementation of copayment, prior authorization from a state medical consultant for more than two ambulatory services or more than two prescriptions in any month was instituted. Although the impact of prior authorization was not quantifiable, there is reason to believe that its effect on utilization patterns was at least as pronounced as that of the copayment provision. Second, the experiment was marred by the short time span in which it was conducted. The experience of both copayers and non-copayers was examined for three months, which makes it impossible to analyze the long-term effects of copayment, especially those involving erosion of deterrent aspects. Finally, in a functional sense, the experiment was conducted with no genuine control group. The group used for comparison differed from the copayer in that it had lower income and/or lower resource levels, yet income, it has been frequently shown, is a key variable affecting utilization patterns. In this context, the comparative results provided by the Medi-Cal experiment are dubious, at best.

Recent evidence provided by Roemer et al.[7] suggests that the deterrence effects of copayments may have adverse effects on patient health, resulting in increased hospitalization. The study examined data from the California copayment experiment and found evidence that, following introduction of prior authorization for physician visits and prescription drugs, utilization of physician services and prescription drugs fell for both copayers and non-copayers. The decrease in utilization was greater for the former group. The reduction in utilization rates remained below previous (to prior authorization) levels throughout the first year of copayment for both groups. The differential between copayers and non-copayers remained virtually the same as in the quarter in which prior authorization was introduced. The authors infer that the differential was due to copayment, a conclusion which seems unwarranted. From the data, one simply cannot disentangle the effect of prior authorization and copayment.

The Roemer study, then, provides data that, in the first three quarters that copayment was in effect, hospitalization rates were higher for copayers. They conclude that increased hospitalization is related to the deterrent effect of copayment on ambulatory services:

> When such ambulatory services are inhibited, it would seem that a price is paid--namely, a rise in the relative rate of hospitalization. It is likely that this elevated hospitalization rate is due to the postponement of ambulatory care, so that when the patient is finally driven to seek assistance, his case is more advanced and requires inpatient care.[8]

This conclusion seems extremely strong, given the evidence provided. It is based on an examination of three quarterly averages of hospitalization rates; the fact that the hospitalization rate of the copayers in the fourth quarter of the copayment experiment fell well below that of the non-copayers is ignored. A small proportion (2 to 3 percent) of each group were hospitalized in each quarter; thus, quarterly averages are acutely affected by small random variations. Moreover, hospitalization rates for both groups were 10 to 15 percent below those for the period when prior authorization was introduced, which argues against substantial increases in untreated illness. Finally, in a subsequent paper,[9] Hopkins, Roemer and colleagues discuss results from a survey of California Medicaid providers and conclude that prior authorization was a much greater problem for them than copayment. It is therefore very difficult to attribute any utilization changes to Medi-Cal's institution of copayment when institution of prior authorization was such an important added factor.

Almost without exception, the main focus of copayment experiments conducted to date has been on the demand side; that is, the analysis has emphasized the role of and reactions by the patient/consumer in a copayment scheme. Although this information has been useful and interesting, a far more important dimension of medical utilization has gone largely unexamined; namely, the supply (i.e., the provider) response to copayment. Arguments that cost sharing will be successful in controlling costs assume that the copayments will be collected by physicians and that consumers will adjust their level of utilization in some way. However, with an extremely low-income population such as Medicaid, the copayment charges will frequently be difficult to collect.[10] Physicians will suffer losses of income by the amount of uncollected copayments.

The issue of provider response is the same as that raised in the physician reimbursement paper of this series.[11] That is, there are several possible supply responses, and our knowledge of physician behavior is such that the outcome is difficult to predict. Physicians may respond by not participating as heavily in the program, by seeing fewer patients, or by providing fewer services of, perhaps, lower quality. Several state administrators informed us that physician reimbursement levels were already very low relative to private charges, and that the burden of collecting copayments would only further alienate physicians. Reductions in private physician participation can result in increased use of hospital outpatient departments or emergency rooms in which the costs of visits to the state are higher. Alternately, because of physicians' ability to influence demand, e.g., initiating follow-up visits, referrals, hospitalization, lab tests, etc., they may respond by prescribing greater amounts of discretionary services in order to attain their income targets. Evans[12] has persuasively argued that, while cost sharing may constrain patient-initiated demands, it will not affect overall utilization rates because physicians will respond to reduced patient-initiated demand by demand creation.

Copayments may also result in changes in the billing behavior of physicians. Copayments which are fixed in dollar units (e.g., two dollars per visit) invite physicians to bill for more complicated procedures than are actually rendered. For example, the coding system in the 1974 California Relative Value Schedule lists five types of office visits for new patients (brief, limited, intermediate, comprehensive and unusually complex) and seven types for established patients, each with a different value. The physician can ignore collection of the copayment and increase the proportion of billings for procedures of higher value. This situation argues for the use of coinsurance, where patient liability is fixed as a percentage of the price of the procedure, rather than the use of copayments. However, coinsurance liabilities have proven difficult for physicians to calculate and for patients to understand and, as a result, collection problems have been substantial. New York eliminated its coinsurance liability on physician services in 1974 because of widespread physician opposition.[13]

A final option for physicians is to not bother with collection of copayment charges but, instead, to raise fees. Clearly, physicians cannot do this if the state maintains a fee schedule. However, if the state uses a customary and prevailing charge reimbursement system, with annual updating, actual charges will eventually be reflected in higher customary and prevailing charge ceilings. As a result, the burden of copayments can, over time, be passed back to the state.

C. Hospitals

States are permitted to impose nominal levels of cost sharing on the medically needy for hospital inpatient care but, technically, they may not impose charges on cash assistance groups. In the fall of 1975, Georgia received a waiver from HEW and imposed a deductible of one-half of the cost of the first day of hospital care or $25.00, whichever is less, on all recipients. On June 14, 1976, a federal court ruled that the Georgia Medicaid agency was permanently enjoined from continuing the copayment project unless it complied with federal regulations protecting human subjects in research. As a result, the deductible is no longer in effect. Georgia is the only state which had adopted any form of cost sharing for hospital inpatient care in its Medicaid program. Nonetheless, the rationale behind the effort is worth examining.

In theory, a deductible will deter unnecessary admissions because physicians will be conscious of the burden to be borne by the patient. When technically feasible, the physician will provide the same services in an office or clinic. As a result, Medicaid costs will fall because of fewer admissions and because services delivered on an ambulatory basis are less expensive than the same services rendered in a hospital. Unfortunately, the effect may be that necessary admissions are sometimes delayed, with serious effects on patient health. In addition, unlike copayments for physician services and prescription drugs, a hospital deductible will be imposed on a relatively small share of the Medicaid population and, in many cases, at a time of serious illness. Particularly in states with low levels of welfare benefits, even small hospital deductibles can be a large burden for poor families.

The first question which arises is whether the hospitals will assess the charge. If the deductible is assessed, patients will either pay it with their own or with relative's resources, or their admission will be deterred. If patients are not deterred and do bear the charge, the deductible amounts to a particularly regressive method of financing Medicaid hospital costs. Not only is it doubtful that the Georgia provision would have been politically popular; it would have generated less than $2.5 million or less than 1 percent of the state Medicaid budget, assuming no deterrence.

Unfortunately, there is less evidence of the impact of small deductibles on hospital admissions when imposed on low-income families than there is evidence of their impact on physician services. If hospitals impose the charge, it is likely that some deterrent effect will be observed. The literature on hospital insurance and utilization tends to support the argument that admissions rates will be negatively related to such deductibles and to coinsurance as well.[14] It is noteworthy that the response to deductibles and that to coinsurance may be quite different. Small deductibles may deter some admissions but may encourage patients to remain hospitalized longer and use more services once hospitalized, in part because of the zero price for the "extra" days and services, and in part to maximize benefits from the initial outlay. Coinsurance would probably have less effect on admissions but may reduce the length of stay and the number of ancillary services provided the patient. If it has such effects, it may be a more valuable policy instrument than deductibles, despite the more burdensome administrative tasks given the hospital. Coinsurance would also impose an added administrative burden on the state if it chose to maintain records of patient coinsurance liabilities in order to limit the liability of individuals or families.

There are strong reasons for believing that savings to society and to the state will be insignificant, even if coinsurance or deductibles are collected and do have some deterrent effect. If admission rates fall because of a deductible, or if patient days and ancillary services decline because of coinsurance, occupancy rates and hospital revenues will fall, with little reduction in the short run in hospital costs. However, hospitals can recoup lost funds in a variety of ways, some of which will affect Medicaid. Medicaid, as described more fully in the hospital reimbursement paper of this series,[15] pays on a reasonable costs basis. In computing Medicaid's share of a hospital's costs, total hospital costs are multiplied by the ratio of charges made to Medicaid patients to charges made to all patients. If charges made to Medicaid patients fall because of the deductible, the share of the hospital's costs borne by Medicare, Blue Cross and all other cost-based reimbursors will increase. Increasing charges on all services to cover costs will shift the burden to the commercially insured who pay on a charge basis and, in some cases, to the uninsured. Deficits of county and municipal hospitals will be the burden of local taxpayers. However, hospitals can also limit loss of income from Medicaid itself by increasing their charges for services used disproportionately by Medicaid patients or decreasing charges used disproportionately by non-Medicaid patients. The way in which costs will be passed on to the various carriers will likely depend on the hospital's patient mix. It is clear, however, that through accounting manipulations and changes in their charge structure, hospitals can avoid significant losses. Furthermore, they can, if they choose, pass much of the apparent Medicaid savings back to the program.

It seems even more likely, however, that hospitals will not attempt to collect the charge and will continue to serve as many Medicaid patients as before. This will be particularly true of hospitals for whom Medicaid patients are a substantial share of clients. With an average expenditure of $685.00 per hospitalized Medicaid patient, it seems unlikely that the hospital will be greatly concerned with collection of a $25.00 deductible. The last days of a hospital stay are relatively low in cost. As a result, it also seems that it will not be difficult for hospitals to recoup lost income by extending hospital stays. Evidence on changes in case-mix as occupancy rates fluctuate suggest that hospital administrators have enough control over hospital operations to affect lengths of stay.[16] Even if a state maintains a sophisticated utilization review system and examines lengths of stay against diagnosis-specific criteria, it will not be difficult for hospitals to extend enough patient stays to recoup a substantial amount of lost income.

D. Prescription Drugs

Several states introduced copayments for prescription drugs in 1975. As of August 1976, Alabama, Arkansas, Georgia, Kansas, Maryland, Mississippi, Montana, North Carolina, South Dakota and Virginia all imposed some form of copayments for prescription drugs.[17] One reason is that prescription drugs are an optional service; copayments can be imposed on both the categorically needy and the medically needy, thus eliminating the burden on providers of determining whether or not a Medicaid recipient is or is not liable for copayments.

The rationale behind cost sharing for prescription drugs rests with the argument that there is considerable overutilization of prescription drugs in the United States.[18] Lee and Silverman cite several factors for the massive use of drugs, finding that the public, the physicians, and the drug industry are all at fault.

> Much of the use of pharmaceutical agents in this country has been rational and brilliantly successful. Lives have been saved, pain has been alleviated, the anguish of severe mental illness has been tempered, crippling has been prevented, recovery has been made quicker, and some age-old plagues have been essentially obliterated. But it is only too evident that a significant portion of drug use, both prescription and over-the-counter, is unnecessary and irrational. This misuse is largely the result of (1) the lack of skepticism or the outright naivete of both physicians and patients; (2) the flood of advertising and other drug information that has too often been unobjective, incomplete, or misleading; (3) the lack of readily available drug information that is objective; (4) the widespread use of a drug on a usually false premise that "it may not help but it won't hurt"; and (5) the equally widespread belief among the public that there is or must be a pill for every ill.[19]

There seems little argument that there is an excessive amount of drug prescribing in the Medicaid program. A recent study[20] of the use of Medicaid services in California found, after controlling for demographic characteristics and other variables, that prescription drug rates in urban counties increased significantly with the availability of physicians in a county and the proportion of physicians who were general practitioners. While this is not evidence of over-prescribing, it does suggest that factors other than medical necessity affect the use of prescription drugs.

Copayments would provide some limit to excessive provision by imposing a monetary penalty on Medicaid recipients.

It can be argued that the prescribing of drugs is done by physicians and, thus, that it is unfair to impose charges on patients who are largely ignorant of the risks involved in failure to follow the physician's instructions. On the other hand, it has been argued that physicians often prescribe drugs when time is limited, diagnosis is uncertain, and patients will be dissatisfied if no remedy is offered. A prescription provides evidence that a solution to their miseries is in progress. In testimony before a House Subcommittee in 1972, Dr. Charlotte Muller has stated:

> [T]he typical physician can spend an average of only 17 minutes with each patient. At the end of that time, even though a patient's problem may not be completely discussed, understood or solved, it is necessary for the physician to turn his attention to the problems of the next patient who is sitting in the waiting room. But the physician must first get rid of the patient then in this office, and in a manner which will make all parties comfortable. Prescribing is, theoretically at least, a means of terminating the interview in a fashion that satisfies both doctor and patient. The prescription is a signal for the approaching end of the encounter, it both summarizes and carried forward the relationship, it is an expression of concern, and it deals with the interests of both parties in a manner perceived as equitable.[21]

Copayments may mitigate the pressure placed on physicians by patients, reducing the number of prescriptions and Medicaid costs and, perhaps, enhancing patient health.

The limited amount of evidence available suggests that copayments would have a deterrent effect on utilization. Phelps and Newhouse[22] provide evidence from prescription drug studies, using data from Windsor, Ontario, and from the British National Health Services, that the number of drug prescriptions is negatively related to price. The estimated price elasticities in both cases were about .10, indicating that a 10 percent increase in the price facing the recipient would result in a 1 percent reduction in the number of prescriptions.

The California cost sharing experiment also included copayments of fifty cents for each of the first two prescriptions each month. The number of prescriptions fell as one would predict, and this finding was the most consistent of the California experiment. After copayment had been in effect for three months, the number of filled prescriptions fell by 10.5 percent for the AFDC population, 7.6 percent for the aged, and 6.7 percent for the blind and disabled categories. Significantly, little or no differences were found between copayers and non-copayers in the utilization of those drugs viewed to be most important for the treatment of serious diseases. The Roemer et al. study discussed above found that utilization rates for prescription drugs fell for copayers relative to non-copayers after the introduction of copayment, not after the introduction of prior authorization. Both California studies are suspect, however, because of the relatively short follow-up period and other methodological difficulties related primarily to original experiment design.

There are at least four problems with arguments that copayments for drugs will reduce Medicaid program costs. First, the use of copayments for drug prescriptions may lead to increases in injections, when feasible. This practice would increase program costs in those cases, particularly if the program is billed for a physician visit as well. Secondly, the failure to fill prescriptions or to request necessary refills may have severe adverse effects on patient health in certain cases. While most individuals can be expected to react to pain and suffering, not all drugs are prescribed for relief of pain. A particularly important problem may be the preventative and therapeutic needs of children; parents, particularly the poor, are often not good agents or representatives for the interests of their children. Failure to properly adhere to physicians' instructions with respect to drug prescriptions may lead to deterioration of recipient health and increased utilization of physician and hospital services. The net result could be increases in Medicaid costs for these services that outweigh any savings from the initial deterrent effect.

Thirdly, reductions in prescription drug rates may not only reflect the impact of the copayment on patient decisions but also the withdrawal of pharmacies from Medicaid participation. Collection of the copayment may be difficult if there is significant patient resistance. Inability to collect the copayment is, essentially, no different than a reduction in the dispensing fee. Reduction in pharmacy participation will limit the choices of Medicaid recipients, increase their travel time and transportation costs and, because of these costs, lead to fewer filled prescriptions.

Finally, copayments have no effect on the total price of the prescription after the reduction for the fixed charge is made. As a result, patients may request larger prescriptions (i.e., more tablets per prescription) to avoid the copayment for refills. The evidence cited above attests to the effect of copayments on the number of prescriptions. However, if the size of prescriptions increases, the effect of copayment on drug expenditures will be much less. Also, there is no incentive for physicians to prescribe generic-equivalent drugs or for patients to seek lower-priced pharmacies. Coinsurance, where charges are a fixed percentage of the total prescription price, would create the proper incentives but would be administratively more difficult for pharmacists. Coinsurance has proven to be difficult for patients to understand, and collection problems of pharmacists would increase.

E. Conclusions

The concept of "cutting back" in Medicaid implies an unwillingness to pass on the increasing costs of the program to the taxpayers; therefore, cutbacks must affect program recipients. In choosing among cutback alternatives, one should be concerned with the equity and efficiency effects of alternative measures. Equity requires that the burden of program contraction be distributed among recipients in a manner deemed fair, while efficiency dictates that harm to program eligibles be minimized. It can be persuasively argued that cost sharing provides an equitable approach to reducing program costs. With appropriate limits on individual or family liability, cost sharing can allocate the burden of financing the program across all service recipients, without imposing substantial burdens on any particular individual or family. It is considerably more equitable than eliminating optional benefits which, in effect, place 100 percent cost sharing on optional services for the few people in need of them. Cost sharing can be particularly equitable, and can meet efficiency criteria as well, if certain services are excluded from copayments (e.g., services needed by the chronically ill, immunizations, well-child care up to one year and maternity care).

One approach to maintaining equity under a cost sharing scheme is to increase cash assistance grants at the same time that copayments are introduced. The state of Michigan proposed such an arrangement and applied to HEW for a Section 1115 demonstration waiver. Under the Michigan plan, Medicaid recipients would have been required to pay one dollar for every ambulatory visit to an independent practitioner and two dollars for out-patient hospital visits. Since the thrust of the plan was to increase cost awareness on the part of recipients, cash grants to all categorically needy recipients would have been increased sufficiently to cover the copayment responsibility of the average recipient. Since the state did not want to disrupt the provision of preventive or obstetric care and since it recognized that chronic care should not be deferred, these services were exempted from copayment requirements. However, since the plan applied a copayment to physicians and out-patient services for the categorically needy, it violated federal regulations and required a federal waiver. The waiver was denied, and the plan was not implemented. Ironically, if Michigan had cut back its physician and out-patient services to one per month, a select group of patients--the chronically ill--would have suffered egregious harm, and preventive and obstetric services would have been all but eliminated, yet the federal government would not have intervened.

The evidence on the efficiency effects of cost sharing is less clear. The limited empirical research on the effects of copayments on the poor leads one to conclude that copayments may have a limiting effect on the utilization of physician services, at least in the short run. Evidence from Canada suggests that the initial effects wear off, perhaps because of changes in patient attitudes or incomes, inflation, or changes in delivery practices of physicians. The evidence from Scitovsky and McCall suggests that the deterrent effect does not wear off. In the long run, the ability of physicians to influence patients' decisions can result in an increase in the supply of certain kinds of services in response to reduced patient-initiated demand. Nonetheless, the short-run effects can be important for a state faced with severe budget problems. Copayments on physician services may have adverse effects on patient health, leading to the increased use of other services. However, the presence of this type of impact has not been substantiated. Finally, a very important result of copayments could be the reduced participation of physicians in private practice, leading to increased use of hospital outpatient facilities at perhaps greater cost to the state.

Cost sharing on hospital stays is unlikely to be used by state Medicaid programs because of legal considerations. If cost sharing were implemented and hospitals did collect the charge from patients, such payments could very easily become a serious burden on patients. However, it does not appear difficult for hospitals to recoup lost funds in a cost-based reimbursement system through manipulation of their charge structure or through extension of patient stays. Thus, any reductions in admissions or lengths of stay will, most likely, fail to reduce the amount of funds allocated to the hospital by the public or the state.

Copayments for prescription drugs, as for physician visits, would probably have some deterrent effect and would spread the burden of program reductions across a large share of Medicaid patients. The practice may have several undesirable effects, including deterrence of necessary drug utilization, reduction in pharmacy participation and increase in the size of prescriptions. While the ideal result would be less unnecessary use of drugs, improved patient health and reduced Medicaid costs, the actual result might be quite different.

A final point in defense of cost sharing is that it may offer a powerful technique for limiting such abuses as the delivery of unnecessary diagnostic services, unnecessary referrals to other physicians within a group of clinic arrangement and billing for services not provided. Medicaid patients presently have no financial incentive to oppose such practices; the entire burden of limiting such practices rests on the administrative machinery of the state. Recent experience suggests that states find it extremely difficult to cope administratively with such provider behavior. Cost sharing may limit considerably the need for elaborate and comprehensive monitoring of physician practices.

CHAPTER II

UTILIZATION CONTROLS

A. Introduction

A utilization control may be defined as any regulatory mechanism which materially affects the behavior of providers and/or patients in terms of the quantity, quality or type of medical service utilized. More specifically, utilization controls are designed to influence utilization patterns in such a way that unnecessary or inappropriate services are minimized for any given benefit package. It is not surprising that utilization control is considered a primary element in Medicaid cost containment policy, for it appears to offer something for nothing. Unlike benefit limitations or restrictions on eligibility, the client need not suffer any real reduction in program benefits if the control mechanism is properly administered. Likewise, a utilization control need not result in lower levels of remuneration for medical providers, at least not for those who meet the program's standards for medical necessity.

This chapter will examine both the limitations and the potential for cost containment of utilization control. The chapter is organized into four major parts. The first delineates the requirements for effective utilization control and provides a checklist of factors a state agency can use in designing new control mechanisms or in evaluating progress under existing programs. The second section analyzes the need for utilization control from the empirical perspective of current estimates of the prevalence and cost of misutilization in medical resources. This section is followed by sections which discuss and analyze all the major control mechanisms presently employed by private and governmental agencies. The statutory limitations and effectiveness of each approach are judged with respect to possible adoption by a state Medicaid agency. The concluding section offers a priority listing of the utilization control options open to the states.

B. Requirements for Effective Utilization Control

Several fundamental requirements must be met if utilization control is to be an effective means of cost containment. First, the problem must be definable in an operational sense. If utilization control is to reduce the incidence of unnecessary or inappropriate medical utilization, then objective criteria for identifying the existence and scope of such practices must be agreed upon. This issue is perhaps the thorniest in the whole cost containment debate: what may appear unnecessary from one perspective is often judged essential from another. Three general types of criteria are at issue here. The question of medical necessity is the first. Is the procedure warranted from a medical standpoint? If so, does the attending physician make the judgment, or is it based either on the usual practices in the community or perhaps on some national standard? The second question relates to medical appropriateness. A procedure may pass the test of necessity but still be judged inappropriate as to setting or intensity of resource requirements. The final question concerns social priorities. Since it is a fact of economic life that the decision to employ resources for one endeavor precludes their use for some other purpose, does the medical procedure represent the best use of society's scarce resources?

When standards of necessary and appropriate medical utilization are defined, care must be taken to ensure that any significant deviation from these standards can be detected. A surveillance system for the collection and interpretation of utilization data is required. The data collection function of such a system may take many forms, including on-site periodic censuses of patient care, claims reviews, or sample surveys. The choice will depend among other factors on statutory requirements, cost and degree of desired accuracy. Regardless of which approach is employed, the data must be capable of interpretation. For this purpose, most systems employ some combination of appraisal norms based on statistical measures of observed performance and/or medical standards designed to establish acceptable limits of variation around the norm for individual cases. The initial purpose of such systems is to measure the magnitude of misutilization and, more specifically, to provide the basis for estimating the net cost attributable to the problem.

If it turns out that this cost is unacceptably high (this is the general presumption behind utilization control), the next requirement is that some form of regulatory action be available to bring the delivery system back within acceptable limits. There are literally dozens of possible mechanisms but only two basic approaches. The first approach is to prevent or at least minimize the probability of misutilization before it arises. Such controls as preadmittance screening and prior authorization of elective hospital admissions, second consultations for surgery, certification and recertification of medical necessity and concurrent review all fall within this category. The second approach is retrospective and is designed to ensure that payments are made only for "necessary" services. The theory behind such mechanisms as institutional utilization review and claims review is that, if providers know they are being actively monitored, they will be less likely to recommend inappropriate or unnecessary treatment.

The choice of which utilization control to employ depends, ultimately, on political considerations which in turn reflect the relative bargaining power of the affected constituent groups (providers, patients and taxpayers). The state agency may have little say in terms of the final outcome, particularly if it takes a passive stance with regard to the available options. In order to maximize its influence, the state agency must take an active role in defining those mechanisms which are most effective given the constraints under which the program operates. To be effective, a utilization control must be (1) appropriately targeted, (2) enforceable and (3) efficiently administered. Since these requirements are frequently violated in practice, the implications of each deserve consideration.

Appropriate targeting means that the control is designed to meet its objective--in this case, cost containment. It requires that the causal factors underlying misutilization be understood well enough to allow the development of tractable solutions for those problems considered sufficiently wasteful to warrant intervention. It also requires that the control mechanism be uniquely targeted. Clearly, cost containment is not the sole objective of a public financing program; if it were, the solution would be to eliminate the program altogether. Cost containment is, in fact, a secondary objective to be met, subject to the realization of the program's primary goals relating to services, clients and providers. Under these circumstances, unique targeting represents a method of constrained optimization in which the constraints are framed in terms of predefined achievement levels for the program's primary goals. A utilization control, for example, can effectively control costs only if acceptable levels of quality are determined beforehand. To expect a single utilization control to simultaneously contain costs and maximize quality not only defies all logic, it makes it impossible to meet either objective.

An effective utilization control must be enforceable. This statement may appear self-evident, but it represents another critical point all too often overlooked. Under our form of government, effective law enforcement requires widespread voluntary compliance which presumes a general acceptance of the objectives and mechanisms of the enforcement procedure. As will be shown later, some forms of utilization control are subverted at this stage because of basic disagreements over the mechanism, if not the intent, of control. However, it should be stressed that a lack of compliance may also be expected whenever providers or patients become convinced either that the control will not be enforced or that the penalties associated with enforcement are trivial. Specific procedures for the control of this latter type of noncompliance should include (1) unambiguous definitions of all prohibited actions, (2) an efficient and publicized system for monitoring compliance, (3) clearly delineated and nontrivial penalties and (4) timely and consistent enforcement. In general, these procedures can be successfully pursued only if the agency responsible for achieving the goal of cost containment has the authority to administer the control and is held accountable if the goal is not met.

The final requirement for an effective utilization control is efficient administration. In the private business sector, "efficiency" is usually measured in terms of profits. According to classic economic theory, profit motivation not only guarantees that costs will be minimized for each level of output but also ensures that no endeavor will be carried out beyond the point where the cost of production exceeds the market price. If one were to employ this criterion in public administration, the objective of eliminating unnecessary or inappropriate utilization would never be fully realized, simply because the extra costs of eliminating the last vestiges of misutilization are bound to be higher than the benefits derived. The public outcry for efficiency in government should be heeded, but it must also be understood that in some cases (e.g., fraudulent billing for services not provided) the public may demand that all such practices be eliminated even if it is "inefficient" to do so. Under these circumstances, the appropriate standard for the public agency is "cost-effectiveness," whereby the scale of operation reflects a balance of economic and political considerations.

C. The Extent and Cost of Unnecessary and Inappropriate Utilization

It should be stressed at the outset that there are no universally accepted standards for unnecessary and inappropriate medical utilization, making it impossible to estimate with any accuracy either the extent of the problem or the economic costs associated with it. However, a public agency charged with cost containment need not operate in the dark, for there is considerable circumstantial evidence that the problem is significant in each of the three categories of misutilization.

1. Medically Unnecessary Services

A medically unnecessary service may be defined as any treatment procedure which could be eliminated altogether without harming the health of the patient. In fact, given some positive probability for iatrogenic complications, the elimination of unnecessary services should both reduce medical care costs and raise general health standards. The problem is that medical protocols are seldom specific enough for the attending physician or a review panel to state unequivocally which procedures are necessary and which are not. Attempts to get around this problem of legitimate differences in professional medical opinion have taken two distinct tracks. The first is site-specific and combines professional peer review with judgmental standards of what constitutes misutilization. The second concentrates on statistical differences in treatment levels across states, regions and delivery systems. Both approaches have centered on hospital care because of the presumption of large potential savings in this area.

A review and critique of the pertinent peer review literature is contained in several recent articles and need not be repeated here.[1] However, one series of site-specific studies of hospital utilization patterns does warrant mention because of the recent attention it has received in Congress and in the press. Organized and conducted by Dr. Eugene McCarthy, the studies analyzed the effects of second consultation for elective surgery on behalf of union employees in New York City between February 1972 and December 1975.[2] According to preliminary estimates, McCarthy found that an average of 24 percent of all surgical procedures initially recommended were not confirmed by the second opinion.[3] Despite the problems of extrapolating these results nationwide,[4] the Subcommittee on Oversight and Investigations of the House Committee on Interstate and Foreign Commerce (the Moss Committee) concluded from the evidence presented that 2.38 million unnecessary surgeries were performed in 1974 at a total cost to the American public of $4 billion.[5] The specific cost to the Medicaid program was set at $272 million.[6]

In a follow-up study, McCarthy traced the records of union members receiving negative indications for surgery on second consultation and found that three-fourths had not had the surgery performed within one to four years after the consultation.[7] Based on these new findings, McCarthy estimated that between 11 and 13 percent of all persons told by their physician to have an operation probably do not require surgery. While well below the Moss Committee estimate of 17 percent unnecessary surgery,[8] these figures, if correct, nevertheless reflect a substantial waste of resources.

Evidence of unnecessary surgery is also posited in the case of a number of recent cross-sectional analyses. For example, Wennberg and Gittleson have shown that the age-specific probability of tonsillectomies ranges from 8 to 62 percent in different areas of Vermont; a difference which they conclude cannot be explained by medical necessity.[9] Similarly, Gornick's finding that surgical procedures account for 34 to 36 percent of all Medicare discharges in the Northeast but only 25 to 27 percent in the South cannot be "explained by variations in the patient characteristics examined."[10] Cross-county as well as cross-national studies have tended to show that the number of surgeries is in proportion to the number of surgeons--and not the medical need for surgery.[11] A recent study of differences in surgical rates under the FEHBP program found that the probability of a subscriber having had a tonsillectomy by 1970 varied from 41.3 percent (Blue Cross) to 10.5 percent (prepaid groups).[12] Other studies of surgical practices in prepaid groups and HMOs have consistently shown lower rates of surgery (and hospital utilization in general) than under traditional financing-delivery systems.[13]

Next to surgery, overlong hospital stays (and, to a lesser extent, overlong nursing home stays) have received the most attention in the literature.[14] Just as a patient may be admitted to a hospital or nursing home unnecessarily, "legitimate" patients may stay beyond the point of medical necessity. Here, too, the circumstantial evidence of misuse of resources is considerable, but, unlike the case of unnecessary surgery, most analysts are hesitant to suggest that extended stays serve no medical function. Rather, the majority opinion appears to be that medical care in such cases could be provided more cheaply in another setting; for this reason, discussion of the seriousness of the problem will be deferred to the next section.

While hospital care is an obvious target for cost containment strategies, there is no a priori reason to suspect that unnecessary utilization is any more prevalent here than in other forms of medical treatment. Overuse of drugs,[15] antibiotics[16] and dental care[17] have all received attention in the recent literature. However, the one area that has received very limited exposure is perhaps the most obvious of all; namely, medically unnecessary physician visits. There are a number of reasons for this apparent oversight. First, a large percentage of all physician contacts are patient-initiated; public programs have, with few exceptions, been hesitant to place any administrative limits on patient access (it is, of course, largely through the imposition of utilization controls that statistics on misutilization are generated). Secondly, one of the expressed purposes of the Medicare and Medicaid programs is to increase physician access on behalf of the poor; program administrators have tended to view the resulting increase in utilization as an indicator of success and not as a potential area of failure. Finally, and perhaps most importantly, the increasing emphasis on preventive health care has led to a general, albeit unwarranted, presumption that the more pretreatment patient-physician contacts the better. Needless to say, the lack of information on a problem is no guarantee that the problem does not exist. Physician resources are wasted in a medical sense through the demands placed on the system by the worried well and by those seeking social contacts rather than medical care. Similarly, it is reasonable to assume that some percentage of physician return visits and specialist consultations are also medically unnecessary. In neither case is the magnitude known, nor is it known whether the problem is greater for Medicaid eligibles as opposed to the general population.

2. Medically Inappropriate Utilization

The notion of medical appropriateness has both a clinical and an economic side. Any treatment setting may be appropriate in a medical sense if it leads to the desired outcome of improved health status. But some settings are more expensive than others, and it is sound economic wisdom to use the least costly alternative if the health outcome is going to be the same.

Unfortunately, because the outcome can never be exactly the same, discussions of appropriateness invariably hinge on some trade-off between medical quality and economic interests. The existence of such trade-offs (even though they may be expressed implicitly rather than explicitly) makes it difficult to define inappropriate care, let alone assess its significance. Lacking a precise definition, the typical rule of thumb is to consider a treatment inappropriate if the medical consequences of a lower-cost alternative are slight, compared with the economic benefits of choosing the alternative.

Attention has focused on three such alternatives: the use of nursing home or home health services instead of extended hospital stays, the performance of minor surgery on an outpatient rather than an inpatient basis, and the conducting of diagnostic tests on an outpatient rather than an inpatient basis. Treating these issues in reverse order, the classic study of inappropriate diagnostic admissions to hospitals was conducted by Anderson and Sheatsley in 1967.[18] The authors examined a representative sample of surgical, medical and diagnostic discharges from Massachusetts hospitals over a 12-month period and found that the percentage of diagnostic admissions not requiring hospitalization was nearly double the rate of inappropriate admissions in either surgery or medicine. While the utilization review procedures promulgated by Blue Cross, Medicare and Medicaid may have reduced the incidence of such admissions (estimated at 23 percent in the Anderson and Sheatsley study) since 1967, it is highly probable that room for improvement still exists.

The idea that many simpler forms of surgery, as well as diagnostic tests, can be effectively performed on an outpatient basis is not new. However, the recent development of "surgicenters" and the continued growth of HMOs have spotlighted the opportunities for substituting cheaper alternatives to inpatient treatment. In his statement before the Moss Committee on national health insurance, Edward Hughes testified that, in one prepaid group practice, nearly 25 percent of all general surgical procedures were performed on an ambulatory basis.[19] Hughes concluded that upward of 20 percent of all surgical treatments now performed in hospitals could be accomplished in outpatient settings, assuming, of course, that such settings are available (which is not the case in most parts of the country). Where outpatient surgical clinics have been established, savings in hospital days are significant. For example, according to one study, such a clinic saved 34,500 hospital days in Phoenix over a three and one-half year period ending in 1974, with commensurate dollar savings to area residents.[20]

Inappropriate length of stay--regardless of admission diagnosis--is generally considered to be the primary source of unnecessary hospital utilization. The evidence (and supporting methodology) is very similar to that presented earlier for surgical treatments. The site-specific peer reviews studies invariably show that hospital stays could be shortened without adversely affecting the health of patients.[21] A study by Zimmer in 1974 exemplifies the typical findings of "significant" bed misutilization found across hospital departments.[22] Zimmer defined misutilization as (1) unnecessary admissions, (2) delays in hospital procedure or (3) delays in discharge. Of the three, delays in discharge accounted for over 82 percent of the inappropriate utilization in the hospital studies. Overstays were directly correlated with the length of the hospitalization episode: 4.9 percent of short stays (1-7 days) were judged inappropriate, but a full 24.6 percent of long stays (60 days plus) were considered to incorporate misutilization of hospital resources. Other hospitals face different problems. In a 1972 study conducted by Paul Gertman, for example, 54 percent of inappropriate hospital days in Baltimore were related to delays in performing or receiving diagnostic test results, and 67 percent of inappropriate days occurred in cases involving lengths of stay of 12 days or less.[23]

The cross-sectional studies are even more striking. A report prepared by Clifton Gaus for the Social Security Administration early in 1976 found "inexplicable" differences in Medicare length of stay across 65 conditional PSRO areas.[24] Diagnostic-specific lengths of stay (for an average of 15 diagnostic groups studied) were found to vary from 7.2 days in two California PSRO areas to 13.8 days in Kings County, New York. In the case of surgical admissions, pre-operative days varied from an average of 1.9 days in four Western PSRO areas to 3.9 days in two Eastern PSROs. Postoperative days varied from 6.4 in San Joaquin to 10.5 days in Kings County. Such East-West differences in length of stay are corroborated by other studies including the Gornick report, discussed previously, in which Medicare discharges for appendectomies showed a mean length of stay of 11.2 days in the West compared to 16.1 days in the Northeast.[25] While care must be taken in interpreting these and similar results,[26] it is clear to most observers that such differences indicate a wide latitude in regional definitions of medical appropriateness. If the definitions were standardized, some net reduction in average length of stay could be expected.

Interregional variations in hospital utilization are minor compared to nursing home use differentials. As in the case of surgery and hospital utilization in general, the actual use of skilled nursing facilities (SNFs) and intermediate care facilities (ICFs) appears to be more highly related to differences in the quantity of supply than to any objective measure of need. Research into these differences and their underlying causes has been woefully neglected in the literature to date. But recent evidence presented by the Urban Institute suggests that they cannot be ignored. In Wisconsin, for example, there are currently 101.9 SNF and ICF beds per thousand elderly persons, with a corresponding residency rate of 81.6 per thousand, compared, respectively, to rates of 21.5 and 19.6 per thousand in West Virginia.[27] It is hard to imagine that these differences of nearly five-to-one can be explained by objective variations in medical need between the two states. In part, the low ratio of nursing home supply in West Virginia and other states such as South Carolina, Virginia and Arizona may reflect a conscious decision to ration these services among persons with varying degrees of medical need. But it is just as reasonable to assume that states with high supply ratios are providing a form of housing subsidy essentially unrelated to the necessity for medical care. Direct housing subsidies plus supervised home health care and homemaker services may well represent a less expensive and equally efficacious alternative for many current nursing home residents.

3. Socially Inappropriate Utilization

A third category of misutilization is the commitment of medical resources beyond the point where costs exceed the benefits of treatment. The value of health and, indeed, of life itself may be inestimable in individual cases; yet society as a whole is forced to make choices as to how much of its scarce resources to allocate for medical care as if these values were known. Since the use of resources for medical services precludes employing these same resources elsewhere, it follows that, at some point, further expenditures will not produce benefits commensurate with costs. The question, of course, is how to determine when that point has been reached. Is it 8.6 percent of GNP, as is currently the case? 10 percent, as some have projected for 1980? Or higher yet? Clearly, no easy solution is at hand, but, just as clearly, it is a problem faced each time a state legislature decides to expand or contract its Medicaid program. It is from this perspective that the allocation question is germane to the issue of utilization control.

Cutting medically unnecessary procedures represents the simplest case, since the only costs incurred are administrative. Similarly, the reduction of inappropriate medical procedures is relatively unambiguous, since the cost in reduced medical "quality" is presumed to be slight. However, what happens when we enter what Havighurst and Blumstein describe as the "quality/cost no man's land,"[28] or more precisely, the area in which limiting utilization may prove cost-effective but which nonetheless may result in a measurable decrease in the quality of care? Should case-specific utilization controls be employed in these instances? Logic suggests that they should because the sum of the quality losses arising from such trade-offs could presumably be minimized only if each case were judged individually. However, the administrative cost of this approach, to say nothing of the problems associated with its enforcement, precludes its implementation.

Even if the curtailment of socially inappropriate utilization could be handled more readily through other cost containment mechanisms (specifically, categorical benefit limitations and patient cost sharing), there is still the possibility that inadequately designed utilization controls may themselves contribute to a misallocation of resources. In fact, Havighurst and Blumstein argue that exactly this will happen when PSROs (Professional Standards Review Organizations) begin to grapple with the "quality imperative."[29] For this reason, it is important that the notion of socially inappropriate utilization be kept in mind, even though the agency charged with utilization control may not have the administrative jurisdiction to reduce its impact in a direct sense.

4. Fraud and Abuse

Because utilization control represents a means of monitoring and regulating patterns of usage and provision of medical services, it has the potential not only to curtail unnecessary and inappropriate care, but also to reduce certain classes of fraudulent medical practices. The existence of fraud and abuse in the Medicaid program has been the topic of considerable Congressional and state debate in recent years, with most of the attention focused on the problem of ineligible recipients. However, fraudulent treatment and billing practices represent an equally important area of public concern. This type of fraud may take one of two forms: the provision of services which the provider knows to be medically unnecessary and the billing for procedures not actually provided (or double billing for the same procedure). The first type is probably the more prevalent, but because it is nearly impossible to prove motivation other than "treatment in the best interests of the patient," these cases tend to get mixed in with those arising from honest differences in professional opinion. The second type is more easily identifiable because it is only necessary to check with the patient to see whether the billed service was actually rendered. Even in these instances, however, a provider who bills the state for a few randomly placed "non-services" will, in all probability, never be caught, simply because most detection systems are designed to catch only the large-scale and consistent violator.

For these reasons, it is difficult to estimate the actual extent of fraudulent utilization practices in any particular Medicaid program. Not surprisingly, most of the published evidence is based on sporadic but "spectacular" cases involving, for example, the recent nursing home scandals in New York City,[30] the recent discovery of fraudulent laboratory and clinic practices in Illinois and Michigan[31] and the periodic federal audit reports which show a lack of proper control mechanisms in various state Medicaid programs.[32] It should again be emphasized, however, that a problem cannot be assumed to be insignificant simply because current systems of empirical verification are inadequate to the task.

5. Measuring the Cost of Misutilization

An accurate assessment of the economic cost of misutilization and abuse under the Medicaid program requires knowledge of both the number of cases occurring at a particular time and the net savings to be derived from each case eliminated. As mentioned repeatedly, the first requirement can seldom be fulfilled. However, even circumstantial evidence of misutilization is useful in estimating costs, provided that some minimum analytical criteria are met. The first requirement is that the evidence be based on a valid study design including consistent definitions of what constitutes misutilization and what causal factors are associated with its presence or absence. For example, it is not uncommon for utilization rates to drop immediately after the initiation of a control mechanism such as utilization review or second consultation. Can the difference between the old and new rates then be used to estimate the former degree of misutilization? Not necessarily. If the rate differential turns out to be temporary, it could be ascribed to Hawthorne effects rather than to a true change in provider behavior. One classic instance of just such a sequence of events is found in the case of the New Jersey Blue Cross Approval by Individual Diagnosis (AID) program.[33] Alternatively, the rate differential might be due to factors not at all associated with the imposition of the control. The purported reductions in length of stay under the CHAP program in Sacramento[34] have recently been questioned on the methodological grounds that length of stay had been declining steadily before the CHAP program was begun.[35]

If and when the validity of the evidence can be assured, the next step is to test whether the results can be generalized. Virtually all of the site-specific peer review studies of misutilization can be criticized on this basis. There is simply no way to prove that the existence of significant overutilization in one hospital, for example, is evidence of a widespread problem without substantial corroborating results from other case studies. Moreover, since the criteria for judging misutilization tend to vary widely from case study to case study, extreme caution must be taken in generalizing what appear to be unanimous results. Cross-sectional analyses of utilization patterns avoid this problem but are beset with another of equal magnitude; namely, the difficulty of establishing unambiguous statistical norms for appropriate utilization. In both cases, results obtained for one population group may not be appropriate for another with different characteristics. Even if utilization patterns are correlated to age and primary diagnosis, there may well be justifiable reasons for setting different standards for what constitutes misutilization among Medicaid recipients, as opposed to misutilization among other more affluent populations.

Nevertheless, attempts to measure potential savings in Medicaid utilization are of some value in public policy decisions. At the very least, pilot studies can target the areas where analysis is most needed. When it comes to actual calculations of dollar savings, however, a final methodological point becomes important. Any savings attributed to misutilization must be realizable: administrative and other expenses associated with a control program must be subtracted from gross estimates of the cost of misutilization before actual savings can be calculated. In addition, any estimate of net savings must be based on the marginal, not the average, costs of misutilization. Two examples should serve to clarify this latter point.

If, for instance, misutilization takes the form of surgical or diagnostic procedures performed on an inpatient rather than an outpatient basis, the true cost of misutilization is the difference between the savings realized in reduced hospitalization and the extra costs that would be incurred in treating the patients in an outpatient setting. If appropriate outpatient settings are not available, then, clearly, no savings can be realized. The Moss Committee's estimate that $150 million in

Medicaid funds could be saved by performing approximately 500,000 surgeries per year on an outpatient basis is highly problematical because it is doubtful that current outpatient surgical facilities could handle this extra load of cases.[36]

The second example concerns savings attributable to reductions in unnecessary (as opposed to inappropriate) hospitalization. Again, the Moss Committee findings present a useful target. Although the elimination of unnecessary surgeries may not require significant extra medical treatments in lieu of surgery (an assumption which in itself is questionable),[37] it is still necessary to restrict the cost estimate to realizable savings in reduced hospitalization. Not only did the Committee use hospital cost statistics for New York State (where costs are among the highest in the nation) to generalize their estimate of nationwide savings; they also based their calculations on the average cost per patient day.[38] By definition, the average cost per day is determined by dividing total hospital costs by the number of patient days. The resulting figure includes a prorated amount for all cost categories, including depreciation, interest, capital equipment and salaried labor. None of these costs will be reduced if a smaller number of patients is admitted. In fact, because 60 to 80 percent of all hospitals costs are fixed in the short run,[39] the only savings immediately realized are the marginal costs of food, drugs, tests and other treatment-specific items. Therefore, the Committee's conclusion that $4 billion could be saved annually suffers from a serious upward bias.

In the long run, of course, all costs are variable. What cannot be realized in a year or two may represent potential future savings, provided that capital and other fixed cost elements contract to serve the lower capacity requirements. The proviso is important, for unless hospitals voluntarily reduce per capita bed capacity, some regulatory mechanisms, such as certificate of need, will be necessary to ensure that future bed capacity is planned to meet only medically necessary cases. One cannot assume that such planning capability exists (it does not), nor can one assume that when and if the capability is developed, it will be properly applied. Moreover, if future hospital bed needs could be allocated precisely, the net savings attributable to reductions in future misutilization would still have to be discounted to present values, in order to accurately reflect the savings potential of current efforts at utilization control. The investment in bed capacity for a general hospital is typically designed to last up to 40 years or more. If one were to assume that the "average" bed had 20 years of useful life left in it, the present value of retiring that bed from service would amount to roughly one-sixth of its capital and associated staffing costs calculated at current rates of interest. In other words, even though the returns to bed retirement may represent significant future savings, they are relatively small in current dollars.

6. Causal Factors in Misutilization

The true cost of misutilization is the savings potential under an optimal utilization control program; in other words, the tractability of the solution is as important as the magnitude of the problem itself. Two basic types of solution are possible. A control system may be designed to eliminate the underlying factors which create misutilization (causal targeting), or it may be designed to reduce the deleterious effects associated with misutilization (symptomatic targeting). In general, causal targeting is considered the superior approach, but symptomatic targeting may prove expeditious in terms of short-range results. In either case, confusion over causality may lead to the imposition of inappropriate control mechanisms and to less-than-optimal returns from the program.

Who is responsible for the misutilization of medical resources? The general assumption is that it is the health practitioner in his role as direct provider of services or as the person who recommends and prescribes treatment regimens. The legal rationale for this assumption (in the case of medical services) is derived from state medical practice acts specifying that only physicians may practice medicine. Because unnecessary or inappropriate medical treatments can only be prescribed by physicians, then, clearly, physicians are responsible for the problem--or so the argument goes. Moreover, it is frequently asserted that physicians are motivated to perform unnecessary treatments because of the financial incentives under fee-for-service payment mechanisms; that they prefer "inappropriate" treatment settings, such as hospitals or nursing homes, for their own personal convenience or because in-patient facilities are available to physicians without charge; and that they possess a monotechnic view of medical care which leads to socially inappropriate treatments. It is further alleged that physicians are quick to pick up new medical technologies which are resource using (e.g., CAT scanners) but slow to incorporate technologies which are resource saving (shortened postsurgical hospital stays, for example).

While there is, undoubtedly, some truth to these allegations, they should not be overstated. Misutilization can and does arise from patient pressure for treatment that physicians would not otherwise recommend. Patients may be motivated to demand inappropriate services for a variety of reasons, including a desire for companionship, misinformation on the efficacy of treatment procedures, unstable home environments or simply the fact that insurance coverage makes treatment "free" at the time of delivery. In such cases, a utilization control focused on physician behavior would represent an example of symptomatic targeting, albeit one which might well be effective given physician control over the treatment process.

Symptomatic targeting on the physician will not be effective, however, if the misutilization rests outside the physician's zone of authority and immediate influence. Such is the case whenever misutilization is the product of institutional or health system inefficiencies. The most obvious example of long-run inefficiency, hospital overbedding, has been discussed; the problem also arises when, for example, Medicaid or Medicare patients remain hospitalized only because suitable nursing home placements are unavailable. Similarly, short-run inefficiencies in hospital scheduling patterns, staffing ratios and communications procedures can add "unnecessary" days to a patient stay that are beyond the physician's control.

Unfortunately, there are no reliable statistics on the proportion of misutilization attributable to these causes.[40] The best that can be done is to keep such factors in mind when evaluating specific utilization controls.

D. Classifying Utilization Controls

Although the development of utilization controls represents a recent phenomenon when compared to such traditional regulatory practices as licensure and rate setting, it has surpassed these earlier efforts at control in terms of both variety of approach and regulatory complexity. The complexity makes any classification of utilization controls somewhat arbitrary. The controls could be classified according to target, be it the provider, the patient or the specific medical services. Similarly, they could be classified according to user. The focus in this paper is on the Medicaid program, but it should be obvious that the experience which other users have had with utilization controls can be valuable in designing effective regulatory approaches for state Medicaid programs. It might be noted in this regard that a plethora of utilization controls are employed by other government programs (Medicare, FEHBP, CHAMPUS, the VA, the Indian Health Service, state mental health programs, etc.), by private insurance carriers (specifically, Blue Cross), by provider groups (HMOs, hospitals, medical foundations), by industrial purchasers of health services and by unions.

The most common method of classification, however, is according to the timing of the control relative to the delivery of service--in other words, whether the control is applied before, during or after treatment. In subsequent sections, the control mechanisms falling within each of these three categories will be analyzed with respect to the following schema: (1) definition and objective of control, (2) method of standard setting, (3) current regulatory status vis-a-vis the Medicaid program, (4) extensiveness of control by user type, (5) evidence of effectiveness, including targeting, enforcement, and efficiency of operation and (6) implications for adoption and implementation.

E. Predelivery Controls

There are two basic categories of predelivery utilization controls in use today: those which entail a voluntary or mandatory medical review before services will be considered for payment, and those which require a formal authorization in addition to the review procedure. While nomenclatures differ among users, the first category includes preadmission testing (PAT) for nonemergency hospital admissions, prior certification of medical necessity and second consultations for elective surgery. The second type incorporates all formal prior authorization requirements. Both categories of control are predominantly applied to recipients rather than providers. Moreover, all current forms of predelivery utilization control are service-specific. Most are concentrated in the area of inpatient hospitalization, but a few systems cover some combination of outpatient physician services and podiatry, drugs, dental care, optometric services, medical equipment, hearing aids and various forms of specialized therapy.

1. Prior Review Mechanisms

Definition and Objectives

Since prior review mechanisms are service-specific, the objective in each case is to reduce unnecessary or inappropriate utilization of a particular service. Rarely, if ever, is the extent of the desired reduction specified as an explicit objective by agencies employing these controls. It is equally rare to discover any explicit accounting of the quality constraints under which such mechanisms are to operate. How, then, can one judge whether the cost containment objective of prior review has been reached? The answer is simple; it cannot be done, at least not in any formal sense. Informally, of course, one may consider any reduction in misutilization as an indicator of success, provided that the control itself does not increase program costs associated with the quality of medical care.

Medical Standards and Data Requirements

The potential for reduced misutilization under prior review is derived from two factors: the method of setting medical standards for treatment, and, where applicable, the presumed deterrent effect of peer review. Unlike posttreatment review, there is no system yet established for profiling the necessity for care prior to treatment. While such a system may be developed in the future, current programs are forced to rely on medical standards and criteria rather than on statistical norms. Obviously, this lack increases the likelihood of significant variation in the criteria employed. Not only do standards of medical practice vary from region to region; the differences are even more pronounced from practitioner to practitioner. Few systems of prior review make any attempt to control such variation; in fact, preadmission testing, which represents the most common form of prior review, is little more than a technique of outpatient diagnostic testing by the admitting physician himself. The only standards employed are those relating to the timing of admission and, even here, the permissive character of PAT administration gives the physician and patient the freedom to forego testing procedures.

Standard setting under programs of second consultation is equally haphazard. The usual practice is for a second physician to perform an independent review of a case, based on whatever guidelines he sees fit to use. Depending on the system, the patient may choose either the first or second opinion (assuming a reversal of the initial recommendation for treatment); a third opinion may be advised. The value of second consultations has been questioned on precisely these grounds. When uniform guidelines are not employed, how is one to know whether the second opinion is better than the first? Several answers have been proffered. Perhaps the most curious is the assertion that if a panel of consultants is chosen on the basis of high medical qualifications (e.g., board certification), the second opinion will reflect a higher degree of medical accuracy. While this may well be true, one can then question the efficacy of paying for presumably unqualified opinions in the first place! There are other reasons for expecting the rate of misutilization to decline under a system of second consultations. The first is based on the mere existence of a mechanism to reduce misutilization. If the practitioner is convinced that the program administration is serious about the problem, there may be a behavioral change in the direction hypothesized. The second reason has to do with peer pressure, the assumption being that if doctors know that their recommendations are going to be reviewed by a colleague, their recommendations will rarely deviate significantly from usual practices found within the community. To the extent that the "usual practice" discourages misutilization, a reduction in medically inappropriate treatments may be expected. The final reason is based on the financial incentives which practitioners have, under a fee-for-services reimbursement system, for recommending medical care whether it is necessary or not. One need not overstate the strength of this incentive to hypothesize that some reduction in prescribed treatments will occur, *provided* that the consultant is prohibited from performing the treatment.

It should be emphasized that the behavioral change in each case is hypothesized and not an established fact. The observed reductions in misutilization attributed to programs requiring second consultations (which will be discussed below) may be due to only one of the factors, to some combination or to conditions peculiar to the program in question. However, the fact remains that current systems of second consultation, and indeed prior review programs in general, are not based on uniform protocols, standards or norms of medical practice. Hence, the expectation of savings from the implementation of new control programs of this type is problematical.

From a systems standpoint, the lack of uniform standards is both a major strength and a major weakness of prior review. Administrative expediency is the primary strength. The development, implementation and enforcement of uniform standards is a time-consuming and expensive operation, requiring sophisticated systems of data gathering and analysis. Prior review mechanisms can be brought on line quickly, without the necessity of such data systems. In the long run, of course, these initial savings may be overbalanced by the lack of administrative control implicit in current methods of prior review.

Current Regulatory Status

Federal Title XIX regulations do not directly address prior review utilization controls. With the exception of certification/recertification and utilization review requirements, the states--at least theoretically--have considerable latitude in implementing control programs, but only if they:

> Provide such safeguards as may be necessary to assure that eligibility for care and services under the plan will be determined and [provide] such care and services ...in a manner consistent with simplicity of administration and the best interest of the recipients....[41]

A potentially more serious constraint on state activities is found in a strict interpretation of the 1972 Social Security Amendments which mandate that PSROs shall "assume full authority, where found competent, for assurance of the quality, necessity, and appropriateness of care rendered to patients eligible for coverage under Titles XVIII and XIX."[42] More specifically "when a PSRO is carrying out its review responsibilities, there will be no retroactive review [by the state] potentially leading to the denial of payment."[43]

The issue here is twofold: first, what statutory authority do PSROs have in the area of prior review? and second, what authority does the state have in the event that a PSRO is found to be either not "competent" or not "carrying out its review responsibilities?" On the first point, the statute clearly allows for "advance determination of necessity, appropriateness, and economy" in the case of nonemergency hospital admissions.[44] But, with a single exception, PSROs are given the option of whether or not to impose preadmission review (the exception being elective surgery, in which case "certification should be confirmed before surgery is performed").[45] At present, no PSRO is engaged in either prior review or prior authorization of medical necessity.

Even though PSROs may ultimately have supremacy over state agency actions, it would thus appear that a state could impose prior review mechanisms on the basis that PSROs are not currently engaged in this form of control. Moreover, since existing PSROs are all operating under "conditional" status, the Secretary of HEW is required to authorize any extension of PSRO review activities before it is undertaken. The state is also granted authority to review PSRO performance during this conditional period and to report its findings to the Secretary. While it is unclear at this stage whether a negative finding on the part of the state could induce the Secretary to deny authorization for an extension of PSRO activities in the area of prior review, it is a possibility. But, whatever action is taken, the state faces potential litigation from PSROs if it attempts to expand its own review activities in this area.

Extensiveness of Control by User Type

States which have imposed prior review mechanisms under the Medicaid program have invariably tied them to authorization requirements; discussion of these programs is thus deferred to the next section. Among non-governmental users also, prior review is a relatively uncommon form of utilization control. Although preadmission testing (PAT) is a covered benefit under some 56 Blue Cross plans, under Medicare and under several commercial insurance plans, the evidence suggests that relatively few hospitals make any concerted effort to impose PAT requirements on admitting physicians.[46]

At least five Blue Cross plans have experimented with other variations of prior review, including programs in Vermont, Massachusetts, Michigan, New York and Pennsylvania which provide insurance coverage for second, and in some cases third, consultations for elective surgery. All five programs are voluntary and of such recent origin as to preclude experiential evaluation.[47] A sixth Blue Cross pilot project in Michigan, conducted by the Genesee Medical Corporation (GMC) in conjunction with the UAW and the three major auto manufacturers, represents perhaps the most far-reaching private insurance effort in the area of prospective hospital utilization review. Under this program, hospital admissions in the three-county area served by GMC are routinely reviewed by hospital nurses and physicians for both the necessity and propriety of inpatient treatment before admission. Preliminary estimates by GMC indicate a 20 percent reduction in average length of stay among the cases reviewed.[48]

Prior review mechanisms for inpatient hospitalization are also utilized by a number of employee health plans, including, among others, programs sponsored by New York State, New York City, the United Electrical Workers, the Teamster's Union, Proctor and Gamble and the Raytheon Corporation.[47] The most publicized of these plans are the surgical consultation agreements specified in employee health benefit contracts for the New York City United Storeworkers Union and AFSCME District 37. These programs (referred to previously in conjunction with the Moss Committee report) were both organized in 1972. The Storeworkers' contract requires a mandatory second consultation before elective surgical admissions are approved. Each plan member is required to call a union intake worker to schedule a consultation with a board certified surgeon, who in turn reviews each case for medical necessity. Consulting specialists are not allowed to treat the patient so reviewed, but patients are given the choice of whether or not to follow the advice of the consultant (thus distinguishing these controls from prior authorization). The Council 37 program is similar in most respects, but mandatory second consultations were not instituted until July 1976. Prior to this date, the AFSCME plan operated under a system of voluntary surgical consultations analogous to the Blue Cross projects previously mentioned. The union was persuaded to change its policy on the basis of evidence that a mandatory program had the potential of doubling the number of members who receive nonconfirmations and who subsequently do not go on for surgery.[50]

Evidence of Effectiveness

A utilization control must meet certain minimum standards of targeting, enforcement and efficiency of operation before it can be considered effective and deserving of replication. How do prior review mechanisms measure up? The issue of targeting has already been addressed. Although current prior review efforts have been designed primarily to reduce either unnecessary surgery or preoperative length of stay--important targets in themselves--they fail to meet the criterion of unique targeting in that no uniform standards are employed for either the necessity or the quality of care. The controls may succeed in spite of this shortcoming, to the extent that reviewers place a positive and consistent value on reduced misutilization. But it is reasonable to assume that the strength of those values will vary widely from reviewer to reviewer and may well show an absolute decline as the review program ages.

Enforcement has also proved to be a problem with prior review. In fact, the voluntary programs have no enforcement mechanisms at all, other than moral suasion and the financial incentive of insurance coverage for PAT or second consultations. In Des Moines, Iowa, for example, six general hospitals instituted a PAT Program in 1967. After 27 months of experience, only 1,500 patients had been tested under the new procedures, and over half of these patients were from one hospital which had an average occupancy rate exceeding 100 percent.[51] In 1970, PAT coverage was made available to subscribers of Blue Cross of Western Pennsylvania. Two years later only seven of the 98 participating hospitals in the area were employing formalized PAT procedures, and a mere 642 patients were reviewed during this time period.[52] Experience in other parts of the country is similar.

Mandatory programs such as the Storeworkers' and AFSCME plans for second consultation do not compel the patient to follow the advice of the consultant. Experience under the Storeworkers' plan indicates that approximately one in eight patients decide to go through with surgery even though the consultants recommend against it.[53] Mandatory consultation, however, does appear to increase the effectiveness of the approach. Rough calculations by District 37 suggest that the reduction in elective surgery of 4.4 percent under the voluntary program will increase to 9.4 percent under the new mandatory provision.[54]

The issue of enforcement involves providers as well as patients. A successful second opinion program requires that an adequate staff of reviewers be available and that they share the objectives of the approach.

Both factors depend upon physician acceptance of the control mechanism itself. Second consultation is hardly a new idea, nor is it uncommon for physicians to employ such referrals in their normal practices. However, mandatory second consultation is another issue altogether; the physician may infer that he is incapable of discerning when to ask for help from a colleague. It is on this basis that medical societies have opposed the development of programs such as the Storeworkers' plan. The opposition proved unsuccessful in this case,[55] but in other areas of the country where medical societies have a more powerful influence over their members' behavior, implementation of mandatory prior review programs may be more difficult.

Administrative feasibility alone provides no assurance that a plan will be worth undertaking. In the private sector, the minimum standard for cost-effectiveness is that program benefits exceed program costs. From what little empirical evidence is available, it appears that PAT procedures fail this test. To be sure, the cost of tests performed on an ambulatory basis are seldom more expensive than when performed within the hospital. But what of the benefits? Even in those cases where hospitals require some form of preadmission testing for elective surgery, one finds little hard evidence of success in reducing either unnecessary admissions or length of stay. A recent analysis of surgical admissions under the Blue Cross of Western Pennsylvania plan showed that surgery was performed on 99.1 percent of patients undergoing preadmission testing.[56] Moreover, it was discovered that preoperative lengths of stay were not highly correlated with the presence or absence of PAT. Instead, what interhospital differences were found could be explained by admission and surgery scheduling patterns, workloads in ancillary departments and differences in medical practice.[57]

Second consultations, on the other hand, would appear, at least initially, to meet the minimum standard of cost-effectiveness. The McCarthy studies previously discussed indicate a potential return of 8-to-1, based on an average cost per consultation of between $25 to $55 (depending on the year) and estimated savings of $2.4 million in surgery not performed.[58] Even allowing for such methodological errors as computing savings on the basis of average rather than marginal hospital costs, the savings seem substantial. Is the approach efficient? Probably not to the degree that proponents would like to believe. In the first place, all elective surgery was reviewed prior to treatment, including those cases where, on a priori grounds, one would expect a low rate of nonconfirmations. It is likely, therefore, that the costs of reviewing some procedure types exceeded the benefits obtained. For example, if the marginal savings in not performing a certain procedure are $550 and the cost of review is $55 per case, the probability of nonconfirmation would have to exceed 10 percent in order that the review process be cost-effective. Had the program concentrated on procedures with high rates of nonconfirmation (estimated, for example, at 44 percent for hysterectomy, 40 percent for dilatation and curettage, and 39 percent for prostate surgery),[59] the return ratio would have been considerably higher.

A second point to keep in mind is that an approach which is cost-effective in the beginning may not continue to be cost-effective over time. This point is particularly true for programs designed to change established behavior patterns. One would expect a "successful" prior review technique to reduce both the number of nonconfirmed procedures actually treated and the general level of nonconfirmations as primary physicians adjust to the review mechanisms. But, as the number of nonconfirmations drop, the return ratio also declines until, ultimately, total costs may exceed total benefits. Since none of the experiments with prior review to date have come to grips with this problem, the actual point at which the control becomes ineffective is unknown. If behavioral patterns are highly resistant to change, the control might remain effective for years, but if patterns change quickly, prior review should be replaced with less expensive mechanisms designed simply to ensure that the old modes of practice do not reappear.

The third and final point regarding program effectiveness relates to the characteristics of both the provider and recipient populations affected by the control mechanisms. In the Storeworkers' and District 37 plans, for example, the employer makes a specified annual cash contribution to the union's health and welfare fund. Since money that is used for medical benefits cannot be used for other programs sponsored by the union, the membership has an incentive to make the plan work. In open-ended programs, such as Medicaid, this incentive is not present. Similarly, one would expect prior review to be relatively more effective (but perhaps less necessary) in plans where providers have a financial incentive to keep utilization down (e.g., bonus plans in closed-panel HMOs) than under systems which directly reward service delivery. Because Medicaid operates primarily on a fee-for-service basis, providers are not constrained by this incentive for economy.

Adoption Implications

It should be apparent from the foregoing analysis that much more information is needed before prior review can be recommended as a potential utilization control for state Medicaid programs. In general, however, those circumstances which warrant prior review could be more effectively administered under a system of prior authorization, as will be seen in the discussion which follows.

2. Prior Authorization

Prior authorization and second consultation have much in common; their main difference is in the method of enforcement. Whereas a program of second opinions, even if mandatory, provides the patient with a choice, prior authorization forces the patient to accept the decision of the authorizing agent. There are other differences between the two approaches. Second consultation implies peer review; prior authorization may or may not employ peers in the review process. Current forms of second consultation are appropriate only for services which involve a high degree of medical complexity; prior authorization can be applied to both complex and simple decision protocols. But these differences notwithstanding, both share the same objective and many of the same weaknesses. Standard setting is typically as implicit and case-specific under prior authorization as under second consultation. While authorization requirements may apply to a broader range of targets, neither approach has solved the problem of quality/cost trade-offs in terms of cost containment strategy.

Current Regulatory Status

As in the case of other pretreatment review mechanisms, Title XIX regulations concerning prior authorization are permissive rather than mandatory (subject, of course, to the supremacy clause relating to PSRO activities). On several occasions, the government has attempted to force states to employ prior authorization. In 1970, for example, the Senate Finance Committee recommended that states "adopt procedures for prior independent professional approval of elective surgery, dental care (except for minor procedures), eye care and hearing aids."[60] On November 29, 1974, the Secretary of DHEW went so far as to issue proposed regulations for hospital utilization review which mandated preadmission certification for nonemergency admissions.[61] The response to this proposal from provider organizations and, specifically, from the AMA was so overwhelmingly negative that it was dropped from later regulations.

Current regulations do, however, set guidelines for states employing authorization techniques. Among other cautionary statements, the Medicaid agency is advised that prior authorization must not unduly delay the rendition of needed services, cost more to administer than the value of the service or material to be provided, or interfere with the recipients' free choice of providers.[62] While stopping short of outright prohibition, the regulations also specify that not all services lend themselves to prior authorization. A list of the more common examples where controls might prove appropriate includes placement in skilled or intermediate nursing homes, elective hospital admissions for surgery and diagnostic work-ups, dental services costing more than $35, eyeglasses and hearing aids, physical and occupational therapy, and costly drugs, appliances and medical supplies.[63] The state agency is also directed to systematically study the effects of any prior authorization program from the perspective of improved utilization of care and cost-effectiveness. If the program is not cost-effective, it should be continued only for those providers "whose record exhibits evidence of behavior deviant from community norms."[64]

Extensiveness of Control by User Type

In the private sector, experience with prior authorization controls is limited, for the most part, to a few recent programs designed to reduce misutilization of dental services. The first such effort was pioneered by the Welfare Fund of the United Federation of Teachers in New York City, which instituted a system of predetermination for dental benefits in 1965.[65] Since 1970, several other unions have adopted variations of the UFT plan, including the United Steelworkers of America, the Pennsylvania Public Employees Health and Welfare Funds (AFSCME), and the United Automobile Workers (together with the Big Three auto manufacturers, Michigan Blue Cross/Blue Shield and Aetna).[66] The controls currently in use entail a two-step procedure for predetermination of necessity and, if appropriate, alternative courses of treatment (ACT). Under the UFT plan, all dental procedures costing more than $100 within a period of three months or less must go through the predetermination process. In such cases, the dentist is required to submit his work plan and X-rays to Aetna (the UFT carrier) before payment can be authorized. Aetna in turn conducts a three-tiered pretreatment review process involving claims processors, analysts and, ultimately, an in-house dental consultant for questionable cases. Only the consultant can deny payment authorization, but this decision can be appealed by the dentist, in which case the question goes to a local dental society for peer review.

The predetermination process is supplemented by ACT which, in the case of the UFT plan, was first inaugurated in 1971. The alternative-course-of-treatment approach was implemented largely to give Aetna consultants freedom of action beyond the all-or-nothing decision implicit in predetermination. Not surprisingly, the ACT procedure turns out to be the more useful of the two controls, for it is far easier, on psychological as well as clinical grounds, for practitioners and consultants to reach consensus when there is room for compromise.

In the public sector, prior authorization requirements are found in over half of the state Title XIX programs. The Medi-Cal program in California was the first to adopt this form of utilization control and has gone further than any other state in the extensiveness of its application. Beginning in 1970, Medi-Cal required that all nonemergency hospital episodes be approved by a state medical consultant before admission. A year later, prior authorization was extended to outpatient care and drugs. By 1974, the list of services requiring prior authorization in California included nonemergency hospitalization, extension of stay for both emergency and nonemergency hospitalization, skilled nursing facility (SNF) care, intermediate facility (ICF) care, physician office visits in excess of two per month, pharmaceuticals in excess of two prescriptions per month, all nonformulary drugs, ophthalmic materials, hearing aids, nonemergency medical transportation, durable medical equipment, prosthetic and orthotic devices, home health agency services, dialysis, psychiatric services, and physical, occupational and speech therapy.[67]

The review process is basically the same for all services. When a Medi-Cal client requests a service, the provider submits a Treatment Authorization Request (TAR) which is then reviewed by Department of Health consultants for program coverage, medical necessity and appropriateness, and for certain administrative requirements. The consultant may approve the TAR as is, disapprove it completely or modify the suggested course of treatment. If approved or modified, the provider then renders the service and attaches the TAR to the claims form which is submitted to the relevant Medi-Cal intermediary. The program is large and expensive. In 1974, for example, the Department employed 171.9 professional staff and 329.0 clerical staff in its Field Services Section to process TAR requests.[68]

The consultants employed by the Department reflect the type of service under review: physicians review medical services, pharmacists review prescription items, etc. However, the system deviates from a true peer appraisal process in two respects. First, the initial stages of TAR review are typically handled by clerks, and some TARs will never reach the consultant. Second, for certain services (such as SNF/ICF evaluation of initial admission), the medical reviews are routinely handled by nursing staff.

While no state (with the possible exception of Oregon) has gone as far as California in implementing prior authorization requirements, a number of other states do use similar approaches. These states may be classed according to the extensiveness of their control programs. States which use prior authorization for a relatively large number of Medicaid services include New Hampshire, New Jersey, Maryland, Oregon, Connecticut, New York, Arkansas, Hawaii and Rhode Island. A much larger group requires prior authorization for a few selected services (Colorado, Idaho, Florida, Kansas, Kentucky, Louisiana, Maine, Massachusetts, Mississippi, Nevada, North Carolina, Ohio, Oklahoma, South Carolina, Utah, Vermont, Virginia, Washington and Wisconsin). In Wisconsin, for example, medical consultants are employed to give prior authorization for out-of-state hospital services, private hospital rooms, hearing aids and medical supplies (exceeding $30).[69] Kentucky operates a selective program of prior authorization for identified "overutilizers." Once a Medicaid recipient is so categorized, the case is referred to the "Lock-in" program, a special MA card is issued, and the recipient is limited to four nonemergency physician visits and four pharmacy purchases per month without prior authorization.[70]

Still other states have programs, referred to as prior authorization, which amount to little more than administrative checks rather than utilization controls. Florida's system of phone-in prior approval for Medicaid services is designed solely to prevent ineligible recipients from obtaining services. In Pennsylvania, prior authorization is required for prescriptions and supplies costing more than $15; oxygen equipment and medical equipment costing more than $400 and rental of such equipment exceeding three months; certain medical examinations; and remedial eye care. However, the standards in each case are administrative rather than medical, and if the county assistance offices (which administer the program) have questions concerning medical necessity or appropriateness, their only recourse is to contact the attending physician.[71]

States involved in prior authorization tend to administer their own programs. However, an alternative employed in a few cases involves state contracts with private medical foundations (now PSROs) for performance review functions. The prototype for such contracts is the Certified Hospital Admission Program (CHAP) sponsored by the Medical Care Foundation of the Sacramento County Medical Society. Supported initially by a Blue Shield contract in 1969 to perform claims processing for Title XIX and subsequently by a 1971 EMCRO (Experimental Medical Care Review Organization) grant, CHAP performs both hospital prior authorization and concurrent review for Medi-Cal patients in Sacramento and El Dorado counties.[72] The process of certification requires that the attending physician submit an Initial Request form to the Foundation in advance of admission. This form, containing information on primary diagnosis and proposed treatment, is routinely reviewed by a staff of nurse-coordinators who both approve the requests and establish allowable lengths of stay, based on Foundation norms. The nurse-coordinators are not allowed to deny a request for admission unless the Foundation's medical consultants concur with the coordinator's opinion that hospitalization is unjustified. The nurse-coordinators also perform on-site concurrent review during the course of approved hospital stays. It is this aspect of the CHAP program, rather than pretreatment certification, which has been widely emulated by both EMCRO and subsequent PSRO projects. The one major exception is the Illinois Hospital Admission and Surveillance Program (HASP) which also performs prior authorization but only for certain "non-complaint" physicians and physicians in training programs.[73]

Evidence of Effectiveness

For a control mechanism that has received as much attention in recent years as prior authorization, it is surprising to discover that very few evaluative studies of the approach have been conducted. In fact, there are no published longitudinal studies, so it is impossible to tell whether the primary effect of prior authorization has been to reduce the amount of inappropriate treatments or simply to raise the number of authorizations denied. Without this information, any attempt to ascertain the long-range effectiveness of the approach is hopeless. Even point-in-time studies are deficient. The easiest way to evaluate a prior authorization program is to compare the cost of administration with the dollar "savings" obtained in authorization denials. Aside from the fact that this procedure assumes that all unauthorized treatments would have been carried out in the program's absence, it also presupposes that the control mechanism itself has no effect on provider behavior vis-a-vis the number and type of requests for approval. As discussed in the previous section on pretreatment review techniques, one might expect that a requirement for prior approval would actually induce some reduction in authorization requests. From a policy standpoint, calculations of program savings derived from authorization denials would then tend to understate true savings -- provided, of course, that the estimating procedure is otherwise methodologically correct. The actual degree of downward bias is unknown, but its presence should be kept in mind.

According to estimates by Aetna, the UFT predetermination and ACT programs for dental care save the 250,000 union members and dependents approximately $1 million per year compared to an annual cost of consultations of

$80,000.[74] In March 1976, for example, 250 claims reductions are estimated to have saved $73,000 or about $290 per case.[75] The relatively low cost of dental consultations (currently $3.73 per case reviewed) may be explained partly by the use of clerical intake workers and analysts, but a more important factor is found in the nature of dental practice. Compared to medicine, dentistry is a mechanical operation with relatively established protocols. Where there is professional discretion (e.g., a partial denture for bridge work or extraction in place of full orthodontia), the choice is often determined on economic rather than clinical grounds. However, what makes monetary sense in the short run may be uneconomical in the long run. If, for example, a program of Alternative Course of Treatment (which in the case of the UFT plan is credited with two-thirds of total program cost reductions) creates "savings" by routinely authorizing amalgam fillings instead of gold crowns, these savings may be expected to disappear in later years as patients return for restorative work. (Mention is made of this example not to imply that the UFT program follows such practices, but rather to inject a note of caution for those who might wish to implement an ACT approach.)

Of all the state Medicaid programs employing prior authorization procedures, only the California experience has been extensively analyzed. Based on studies conducted a year after the imposition of hospital precertification, the program was considered responsible in part for a drop in Medi-Cal admission rates of from 17.4 per 100 eligibles in 1969 to 16.1 in 1970.[76] By way of contrast, it might be noted that during the first year of the HASP program in Illinois, the Medicaid caseload increased by 11 percent but hospital expenditures rose by 20 percent.[77] In both cases it is difficult to ascribe the effect to the cause. However, a more recent analysis of prior authorization in California by the Field Services Section of the Department of Health does provide some solid evidence of program effectiveness.[78] Although the reported program costs and savings are flawed by numerous methodological errors, there is little doubt that prior authorization does result in net savings to the state.

The report indicated that all prior authorization programs combined saved Medi-Cal $32 million in TAR denials and modifications in 1974. Given a program cost of $5.8 million in administrative expenses, this savings represents a benefit/cost ratio or rate of return exceeding 5.5 to 1.[79] As noted previously, this estimate is probably understated to the degree that the TAR requirement itself reduced the number of authorization requests. But there are two reasons for suspecting that the estimate is also overstated. First, the savings estimate for TAR denials and reductions in hospital and nursing home utilization was based on average rather than marginal institutional costs. If one assumes that roughly 70 percent of institutional costs are fixed, the true benefit/cost ratio for precertification of hospital admissions drops in the short run from 24.3 to 7.29; for nursing homes, it drops from 6.3 to 1.9.[80] Long-run cost savings properly discounted and based on the assumption of reduced future bed capacity would yield slightly higher estimates. Second, the reported savings were overstated in that some providers who received TAR denials may have gone forward and resubmitted alternative treatment requests. To the extent that such resubmissions were also denied, program savings have been, in effect, double-counted on the benefit side; presumably, the alternative treatment would not have been provided in the program's absence. On the other hand, if the resubmissions were approved, the cost of the alternative should logically have been subtracted from the "savings" attributed to the initial TAR denial. Since this was not done, the results are again overstated. Although the Field Service report does not indicate the number of resubmissions, it is highly unlikely that the net effect of their inclusion would have reduced the benefit/cost ratio for either hospital or nursing home precertification below one. In other words, both forms of prior authorization are probably cost-effective.

The same can be said of prior authorization of physical, occupational and speech therapy and for hearing services, prosthetic and orthotic devices, durable medical equipment, hemodialysis, home field services and psychiatric care--all of which were credited by the Field Service with benefit/cost ratios exceeding two to one.[81] In none of these cases is the fixed cost phenomenon a serious issue, nor is it likely that resubmissions were at a level high enough to push the ratios below unity. Of special interest, however, are the report's findings with regard to physician office visits and pharmacy services. During the study period, prior authorization of physician visits led to a denial of 7,242 TARs for an estimated program savings of $200,000 but at an administrative cost of $953,000.[82] In the case of pharmaceuticals, TAR denials and modifications amounted to 7 percent or slightly over one million authorization requests for an estimated savings of $700,000.[83] The cost of the program exceeded $1 million. The obvious implication in both cases is that prior authorization should be dropped.

The Field Service did advise changes in authorization procedures for drugs and office visits, but it stopped short of calling for a complete abandonment of the approach. Based on indications that most drug TAR denials were nonformulary items, the report recommended that authorization of formulary drugs be dropped, thus saving the state $833,000 in administrative expenses.[84] Similarly, some classes of physicians were found to generate the largest proportions of TAR denials; consequently, the recommendation for prior authorization was retained for psychiatrists, dermatologists, allergists and podiatrists.

Together with other recommended changes in the areas of extensions for emergency hospitalization and SNF/ICF stays and for visual services, the Field Service report indicated that total program savings could be increased as much as $55 million per year, while at the same time reducing administrative costs by $300,000.[85] Even allowing for the fact that 90 percent of the projected additional savings were reductions in institutional care that were miscalculated (at least in terms of short-run savings) on the basis of average rather than marginal costs, the inescapable conclusion is that prior authorization is not only an effective means of cost containment but also that its effectiveness can be increased by selective targeting.

Adoption Implications

California ultimately did drop the prior authorization requirement for most physician services, formulary drugs and certain outpatient clinic services (at the same time, incidentally, that Medi-Cal benefits in these service areas were curtailed). The question remains whether California's experience can be used as a guide for other states contemplating the implementation or extension of prior authorization requirements. The answer will depend in part on the legal and administrative climate within the particular state. For some states it can be expected that PSROs and other professional societies will hamstring any such attempt through legal maneuvering. The cost associated with litigation and delay might well not be worth the trouble. For other states, the sheer complexity of a well designed prior authorization program is beyond the administrative capacity of the state Medicaid agency. One need not go to the same lengths as the General Accounting Office in detailing inadequacies in state Medicaid administration[86] to suggest that states which currently have outmoded claims payment mechanisms will have even more difficulty if claims payment is tied to prior authorization.

By and large, however, most states either have or could develop an effective system of prior authorization. The primary issues are how far to go, which services to include, how selective to be within service categories, and how to avoid needless or capricious delays in authorizing appropriate services. The California experience provides clear evidence that blanket prior approval is not the most effective approach. Generally speaking, prior authorization should not be routinely applied to high-volume, low-value services, such as drugs or outpatient practitioner visits. Even if the per-unit administrative cost of reviewing such services is low, the expected rate of authorization denial is usually well below that necessary to make the approach cost-effective. For example (to generalize from the California experience), if the average price per prescription is $4.81 and the cost per review is $1.03, then the rejection rate would have to exceed

20 percent on a volume of 1,038,000 prescriptions before the program would break even.[87]

This does not mean, of course, that selective prior authorization should be ruled out for such services. One approach used by several states is to allow clients some minimum number of services per month before the authorization requirements apply. Other approaches to selective authorization include cost- or price-specific cut-offs (where only services costing above some minimum amount are prior authorized, as in Wisconsin and Pennsylvania), recipient-specific controls (such as the Kentucky "Lock-in" program), and provider-specific controls (as applied under the HASP program). None of these approaches has been analyzed sufficiently to allow concrete estimates of their true potential, but it stands to reason that program effectiveness will rise in proportion to the fine tuning of the control mechanism. Selectivity makes such fine tuning possible--but only at a price. The California experience suggests that the two-per-month rule for authorization of physician visits and drugs is not effective, but that, perhaps, a limit based on three or four per month could be. Unfortunately, the systems analysis required for a state to answer this question prospectively is beyond the capacity of most state management information systems. Unlike blanket authorization requirements, the application of controls to specified "overutilizers" or "overproviders" means that patient and practitioner profiles must be accurate and up-to-date. Finally, the control tends to lose some of its deterrent effect as the degree of selectivity increases.

Of all the methods, selectivity on the basis of unit cost offers perhaps the most promising general approach. Since all claims payment systems of necessity contain detailed information on service prices and/or costs (depending on the basis of reimbursement), the state has a ready-made source of statistical data on which to set its prior authorization guidelines. Ideally, two types of guidelines should be established: those relating to average reimbursable costs across service categories and those relating to relative cost distributions within service categories. The first type can be used to identify medical services which exhibit the greatest cost containment potential on the basis of high unit reimbursement. Since a state will normally not know the rate of authorization denial before the fact, it makes sense to start with the highest-priced services. Even if the rejection rate turns out to be relatively low, the savings per denial are likely to be substantial. Not surprisingly, if a state were to use such guidelines, the first services targeted for prior authorization would be in the area of inpatient hospital and nursing home care, and possibly in such outpatient services as psychiatric care, hemodialysis, hearing aids and private-duty nursing.

The second set of guidelines for selective targeting within service categories would typically be used only after the state gained experience in the more broadly based service approach. In both cases, targeting on the basis of price provides an advantage not found in other forms of utilization control. By concentrating on price, the state is forced to consider what it (and the client) receives in return. Standards of service quality are thus a necessary complement to price-related guidelines. Those standards may be rather rough in the beginning,[88] but at least the state is set for an explicit accounting of the cost/quality trade-offs in utilization control policy.

One final note on the application of prior authorization to state Medicaid programs is in order. Of all forms of utilization control, prior authorization is undoubtedly the most unpopular with practitioners and institutions alike.[89] Three principal factors are at issue: (1) the control subverts professional prerogatives with respect to treatment options; (2) it creates--by definition--delays in treatment; and (3) it is expensive in terms of the professional time spent in preparing application forms and in reviewing medical histories a second time in the case of application approvals. None of these "costs" will show up in an accounting of state administrative expenses, but they are real nonetheless, and must be considered before a state commits itself to a prior approval program. One might suppose that negative provider reaction would create a problem of enforcement, but this is not the case where payment depends upon prior approval. Ironically, a failure on the part of the provider to comply with authorization requirements may lead to program savings if the provider's alternative is to render the service free-of-charge. Some evidence of this is found in California.[90] A more serious consequence (at least for program clients) is that prior authorization requirements may induce some providers to drop their Medicaid affiliation. Unfortunately, the relationship between state administrative sanctions and provider participation rates is unknown at this time.

F. Concurrent Review

Concurrent review has received the most widespread attention of any utilization control mechanism, in large part because of the emphasis placed on it in the PSRO and earlier EMCRO programs. But concurrent review antedates both of these control programs and reflects one of the government's first attempts to minimize the level of misutilization under Medicare and Medicaid. The original mechanisms were certification and recertification of medical necessity. While these mechanisms still exist, they have been augmented by elaborate procedures for on-site monitoring, length-of-stay checkpoints, quality assurance mechanisms and reporting requirements.

Definition and Objectives

Unlike prior authorization, concurrent review is only feasible for extended treatment regimens; for all practical purposes, it has been limited to inpatient hospital services. The typical procedure is as follows. The attending physician first admits his/her patient and certifies that inpatient care is medically necessary (thereby meeting the initial certification requirement under Medicaid). Within one to three working days after admission, a hospital-based nurse coordinator reviews the case for level of care and establishes a length of stay (LOS) checkpoint based, usually, on age and primary diagnosis. The attending physician is apprised of the coordinator's decision and, if dissatisfied, may appeal to the hospital's physician advisor (typically a staff doctor employed by the local PSRO). The physician advisor has final authority to overrule either the nurse coordinator or the attending physician. When and if agreement is reached, the case continues up to the day prior to the LOS checkpoint, at which time it is reviewed again by the nurse coordinator. The coordinator may recommend either immediate discharge or an extension, based on discussions with the attending physician (thereby meeting the Medicaid requirements for recertification). Any disagreement at this stage goes again to the hospital's physician advisor for arbitration. The entire process is monitored, either by the hospital's own utilization review committee or by the regional PSRO.

While the process of concurrent review is rather clear-cut, the objective is not. The 1972 PSRO amendments which created the current push toward concurrent review established three sets of criteria.

> These reviews will verify that services provided (1) are medically necessary, (2) meet professionally recognized standards of quality, and (3) are provided in the minimum duration and most economical level of institutional care appropriate to the patient's medical needs.[91]

The emphasis here is clearly on cost containment, but in subsequent interpretations, the DHEW Office of Professional Standards Review shifted the emphasis to improved quality.

> The primary emphasis of the PSRO program is on assuring the quality of medical care. Providing quality care may increase health services for some patients in certain areas and could increase costs under those circumstances.[92]

Since each PSRO is given the authority to develop its own statistical norms and medical criteria, there is no way of knowing which of the two objectives--minimizing costs or assuring quality--a concurrent review system should strive to meet. It is true, of course, that not all quality improvements will lead to higher costs, but on the whole, new technological advances in quality of care can hardly be described as cost saving! Under these circumstances, the hospital (and, in the final analysis, the state) is put in the impossible position of employing one control mechanism to meet two conflicting objectives.

Medical Standards and Data Requirements

Concurrent hospital review programs under PSRO auspices will ultimately be required to develop standards relating to the necessity of admissions, length of stay and medical practices within the institution. At this stage, the only relevant standards relate to length of stay.[93] These standards, referred to as LOS norms, will generally require the initial authorization for length of stay to be set at "the point where 50 percent of patients with similar diagnoses and in the same age groups have usually been discharged."[94] Exceptions are made in the case of diagnoses where the usual duration of stay is two days or less. The PSRO is given the authority to establish these LOS norms for area hospital review committees, based on current regional practices. Several sources of information may be used, including PAS-LOS profiles for the region, MADOC (Medicare Analysis of Days of Care) norms, length of stay norms developed by a regional abstracting service such as HUP, or the PSRO's own LOS profiles. Hospitals are not allowed to use their own LOS experience as the basis for such norms, nor is the PSRO allowed to use national or other regional data unless expressly approved by the National Professional Standards Review Council (NPSRC).[95]

Although the PSRO is required to collect certain readily available baseline data, it is not required nor even encouraged to develop its own data processing capability.

> In general, expenditures of (PSRO) funds for purchases of data or data processing should not be necessary. Such requests should include detailed justification with supporting documentation, and should be made to the PSRO's project officer. These requests will be handled on a case-by-case basis.[96]

The PSRO is essentially placed in the position of requiring hospitals to utilize LOS norms based on data over which it has no control as to either validity or accuracy. The issue of validity is the more obvious since in few, if any, PSRO regions does a single abstracting service cover all of the area hospitals. The PAS system comes closest, but in the Northeast, for example, only 25 percent of the acute-care general hospitals are subscriber institutions. Since PAS subscriptions tend to exclude both the largest and smallest institutions, there is a real possibility that area norms based on PAS-LOS will be seriously biased. Moreover, it is inadvisable to combine data from competing abstracting services because of wide differences in definitions and collection methodologies.

The question of accuracy is equally important. If a norm based on the 50th percentile by primary diagnosis is to be useful, then, clearly, the regional data relating to that diagnosis must be accurate. The abstracting services typically derive diagnostic information from hospital case abstracts which are themselves frequently in error:

> Nationwide surveys by the National Center for Health Statistics have indicated that the minimum national average for errors in the principal diagnosis probably runs from 12 to 15 percent and these are gross errors of major misclassification, rather than simple judgmental questions about whether a patient had a mild myocardial infarction or severe chest pains.[97]

The problems are much more serious when it comes to multiple diagnoses. Rather than provide statistical tabulations showing all possible permutations of primary and secondary diagnoses, PAS provides (as do certain other abstracting services) two breakdowns for each principal diagnosis, one for single and one for multiple diagnoses. As is well known, data on second diagnoses are notoriously inaccurate, but without such information, how can an organization possibly establish meaningful LOS norms for complex cases? It cannot be done.

Over time, other methods will surely allow a more valid and accurate classification of LOS norms than is presently possible. When that time comes, it may be feasible to test hypotheses concerning the impact of such norms on the cost of hospital care. All that one can do at this point is to offer a theoretical expectation that average LOS should decline. The reason for this expectation is found in the nature of the PSRO sanctions (payment denial being the most obvious) that can be levied against any physician who consistently exceeds the norms, in contrast to the lack of any corresponding incentives for physicians who are below the norms to increase their recommended average length of stay. But, while there is some reason to expect average LOS to fall, there are limits to the rate of decline. Because of the requirement that only regional standards be employed, one would expect local variations and current East-West differentials to remain.

In the long run, whatever cost containment potential is found in concurrent review may well evaporate when PSROs establish standards for quality assurance in medical practice. As noted in previous sections, there is a considerable consensus of opinion among nonphysicians that too many services are being provided in hospitals, and that overutilization is a primary cause of escalation in hospital costs. This opinion is probably not held by the majority of physicians; it most assuredly is not held by those who have become involved in analyzing the type of process quality criteria which PSROs will ultimately use as standards of acceptable medical practice. In fact, a substantial number of studies have shown that "suboptimal" amounts of hospital services are rendered to between one-eighth and one-third of all patients admitted.[98] If these estimates are to be believed, the fulfillment of PSRO standards for quality assurance will swamp any cost containment objective.

Current Regulatory Status

Concurrent review of hospital and nursing home services is explicitly required under current Title XIX regulations. In fact, it is required at three different levels. At the first level, each State Medicaid Agency is required to have in effect a "Statewide surveillance and utilization control program that safeguards against unnecessary and inappropriate utilization of care and services available under the plan."[99] The utilization control must incorporate certification of medical necessity, recertification for extended services and continuous review of inpatient utilization. At the second level, each hospital and nursing home must, as a condition of participation, have a written plan for utilization review which includes such concurrent review items as admission certification and a requirement that "an initial extended stay review date must be assigned to each admission based on the individual's condition and the date by which there is a reasonable likelihood that the patient can be discharged."[100] At the third level is the requirement that PSROs must, at a minimum, arrange for concurrent review of care provided to inpatients in short-stay general hospitals.

While there is considerable overlap in these requirements, the primary point is that state plan requirements will be met either if all participating institutions meet the utilization review guidelines or if the Secretary grants a waiver under 45 CFR 250.19(b) for a PSRO to undertake the review procedures. There are two major complications with this second option. Current utilization review regulations require, in effect, that participating institutions commence concurrent review procedures before PSROs are established in all areas. As of September 1976, for example, conditional PSROs were operating in only 93 of the 203 designated PSRO regions, with an additional 28 in various stages of planning.[101] But even when all 203 PSROs are operational, the bulk of the review responsibilities will still be in the hands of the individual institution.

> The PSRO will be required to acknowledge and accept, in whole or in part, an individual hospital's own review of admissions and need for continued care, on a hospital-by-hospital basis, where it has determined that a hospital's "in-house" review is effective.[102]

The actual rate at which PSROs have delegated review functions back to area hospitals has been high. In Region I in Massachusetts, for example, 14 of the 18 hospitals (accounting for 84 percent of all Medicare and Medicaid discharges in the region) conduct all of their own review activities.[103] Under these circumstances, PSROs function more as watchdog agencies than as operational review organizations.

However, this should not be taken to mean that PSROs have no power over the review process. All norms, standards and criteria employed by hospitals must be approved by the PSRO in advance. Similarly, PSROs are required to monitor hospital review experience; if the institution backslides in its effectiveness, the PSRO may step in and conduct the reviews itself. Most importantly, the PSRO is immune from direct state intervention in the establishment of standards or in reimbursement for services approved under PSRO guidelines. While subject to some legal interpretation, it is generally conceded that concurrent review is under the authority of the PSROs and not the state Medicaid agency.

Experience Under Concurrent Review and Evidence of Effectiveness

Considering that concurrent review is mandated for all inpatient services provided under Medicaid, Medicare, and Maternal and Child Health programs, one would expect to find a rather substantial background literature detailing past experience with the approach. Such is not the case. Concurrent review is seldom utilized by private insurers or industrial and union employee health benefit plans. Those few plans which do use the approach (generally through contracts with a local PSRO or medical foundation) are all of recent origin.[104] Public sector experience with concurrent review is restricted to a few independent contracts between medical foundations and Medicaid agencies in such states as Illinois (HASP) and Massachusetts (CHAMP) and to a somewhat larger number of EMCRO demonstration grants.[105] In the future, a deluge of experimental data may be expected from the newly formed PSROs. But until then, the basis for judging the effectiveness of concurrent review is limited.

Of the few EMCRO programs which have been evaluated, the CHAP and OSCHUR (On-Site Concurrent Hospital Utilization Review of the Utah Professional Review Organization) plans have received the most attention. As mentioned previously, CHAP has been conducting concurrent review under the Medi-Cal program since April 1970. The actual review process is very similar to the one outlined at the beginning of this section and needs no reiteration. When the program commenced operation, national PAS-LOS norms were used to establish initial length-of-stay checkpoints. These were later revised downward on the basis of evidence that CHAP lengths-of-stay were 1.9 days less for medical cases and 3.5 days less for surgical cases than the PAS averages.[106] While these differences cannot be attributed solely to CHAP,[107] an analysis of length-of-stay trends among CHAP Medi-Cal recipients showed a drop of some 16.7 percent in patient days from May-December 1969 to May-December 1970.[108] The analysis concluded that "the 15 to 16 percent patient day savings by CHAP...appear to be the minimum possible under PSRO programs once they are fully implemented."[109]

This assertion has since been widely criticized on methodological grounds,[110] but most observers agree that CHAP has had some influence in reducing average length-of-stay. The two primary questions are whether the savings are commensurate with program costs and whether comparable savings can be expected elsewhere. There is a divergence of opinion on both points. First, it must be remembered that CHAP employs both preadmission certification and concurrent review; it is not at all clear that concurrent review is the more cost-effective approach. For example, the CHAP program is estimated by Sayetta to have spent 27.46 cents for every dollar saved in reduced hospitalization, compared to 23.71 cents under prior authorization requirements in effect in Fresno County during the same period of time.[111] While both estimates overstate the true savings attributable to hospital review,[112] they are nonetheless indicative of the relative value of the two approaches.

One might argue that the relevant measure of effectiveness is not average net savings but, rather, total program savings. This same study found, in fact, that the reduction in total patient days was nearly four times as high in Sacramento as in Fresno.[113] The problem with a comparison of this sort is that it assumes identical circumstances in each site prior to the imposition of the control mechanisms. As it turns out, in 1969, Fresno County had a 14 percent lower bed-to-population ratio, a 17 percent lower hospital admissions rate and a 16 percent lower average length of stay than Sacramento County.[114] One possible interpretation of these figures is that the margin of overutilization was simply much higher in Sacramento to begin with. An additional difference between the two sites is also important. The CHAP program represented the first experiment with peer review in Sacramento County, whereas in Fresno the local medical foundation had been conducting peer review prior to 1969. The significance of this difference can be seen in the fact that several prototype PSROs have discovered a steady decline in admission and LOS extension denials as physicians learn what will and what will not be approved.[115]

Thus, even if the CHAP experience could be generalized to other areas (which is doubtful),[116] the overall expectation must be that concurrent review will prove increasingly less cost-effective over time. Cost effectiveness will decline not only as physicians adjust to the new regulatory techniques, but also as the costs of review rise. In 1974, the pre-admission review costs under CHAP were $13 to $14 (including preadmission certification), a figure which is expected to double when full PSRO activities are underway.[117] The Department of Health, Education and Welfare estimated in 1975 that concurrent review costs would average $12 per admission.[118] Others have projected costs in the neighborhood of $16.50.[119] Whichever estimate one chooses to accept, the issue is basically the same as that raised in the case of prior review: at some point, the costs of administration may exceed the benefits obtained.

There are, however, three reasons for expecting that concurrent review of hospital services will fail the cost-effectiveness test at an earlier stage than prior authorization. First, concurrent review is an inefficient method of eliminating unnecessary admissions, both because the initial review takes place anywhere between one and three days after admission[120] and because the nurse coordinator has no direct authority to deny admission certification. The savings that are obtained will thus be limited primarily to reduced LOS. Prior authorization, on the other hand, is equally suited to reducing unnecessary admissions (authorization denials) and inappropriately long stays (authorization modifications). Second, as physicians become accustomed to working on a personal basis with particular nurse coordinators, the arms-length bargaining format typical of prior authorization mechanisms is lost. This factor may actually cut both ways, but given the traditional authority and prestige of medical practitioners, it is far more likely that physicians will co-opt the nurse coordinator's prerogatives than vice versa. The third reason why concurrent review may ultimately prove less cost-effective than prior authorization is cost. Physician reviewers are required in both instances, but the on-site staffing requirements for concurrent review result in a cost of between 100 and 300 percent greater per case reviewed than the cost under prior authorization.[121]

From a cost containment standpoint, concurrent review would not appear to be the utilization control of choice (assuming such a choice were allowed under federal regulations). However, this judgment requires some qualification. There is no doubt that the cost-effectiveness of present forms of concurrent review could be increased through selective targeting, either on particular (presumably high cost) diagnostic cases or on those individual providers who have a penchant for exceeding program norms. Such targeting is conceivable under present PSRO regulations, but only at some undefined future date:

> _Initially_, concurrent certification of elective admission will be performed on _all_ elective admissions unless the PSRO can clearly identify in their review plan diagnoses (or problems) or physicians which do not require such review (emphasis in original).[122]
>
> Over time, as the PSRO performs concurrent admission certification it will identify physicians, diagnoses (or problems), and/or institutions which no longer require admission certification ... When this occurs such physicians, diagnoses (or problems) or institutions would not be subjected to admission certification (although continued stay review could still be performed).[123]

But, even in the long run, it is questionable whether selective concurrent review would prove more cost-effective than alternatives such as selective targeting under prior authorization. On a priori grounds, at least, it appears that prior authorization has the edge.

Some critics have contended that concurrent review is superior to prior authorization given the high level of provider dissatisfaction over the latter course of action. This argument implies, of course, that physicians feel relatively comfortable with the peer and non-peer review aspects of concurrent utilization control. The implication is dubious given the fact that large numbers of physicians still object strenuously to the very concept of concurrent review.[124] Moreover, provider dissatisfaction might well lead to more serious problems of enforcement in the case of concurrent review. As one writer has noted, "physicians hostile to the peer review philosophy can impede the efficient operation of the concurrent review system by refusing to heed reviewer's suggestions, by pursuing frivolous appeals, and by failing to maintain adequate medical charts."[125]

There is one problem that concurrent review is unquestionably better equipped to handle. If the purpose of control is to introduce qualitative changes into hospital practices, then, clearly, concurrent review offers an approach superior to prior authorization. However, unless the quality assurance target of concurrent review is established subject to a specific cost constraint, the approach will produce problems just as serious as if the target of cost containment is not subject to a quality constraint. Since present methods of concurrent review are not constrained in either direction by statute or common usage, one is again faced with the impossibility of meeting two conflicting objectives with a single control mechanism. Until this conundrum is resolved, concurrent review must be approached with considerable skepticism.

Adoption Implications

Concurrent review of hospital admissions is a mandated function of PSROs. There is little the state can do to influence either the application or the limitation of this particular utilization control unless local PSROs have yet to be designated. It is true that the state Medicaid agency has a basic responsibility for seeing that providers are in substantial compliance with Medicaid conditions of participation, but the state's leverage in this regard is minimal. In fact, the only realistic option left to the state is in the area of monitoring program effectiveness. The development and implementation of such systems as Utilization and Surveillance Review (S/UR) and the Management and Administrative Reporting Subsystem (MARS) provide an excellent opportunity for ongoing and systematic evaluation of concurrent review activities. If the cost-effectiveness of concurrent review proves short-lived (as is likely), then, at least, there will be hard evidence on which to support a change in present federal regulations.

G. Post-Treatment Review

Post-treatment review differs fundamentally from the other types of utilization control discussed thus far, in that the control mechanism comes into play only after treatment is rendered. There are three generic categories of post-treatment review. Without implying any particular order of priority, they are: (1) fiscal or management review, (2) quality assurance review and (3) utilization review. The first is designed to provide management with information on program effectiveness by department or cost center, both for internal administrative operations and for such external transactions as the status and appropriateness of provider payments. It is found to one degree or another in all systems of claims review and in the MARS and Claims Processing subsystems of the Medicaid Management Information Systems (MMIS). The second type of review is designed specifically to monitor the quality of medical practice and finds expression in such programs as medical audit, tissue review, medical care evaluation studies and quality assurance programs (QAP). The final review function encompasses a variety of techniques for monitoring utilization patterns, including institutional utilization review committee procedures, PSRO profile analysis and the S/UR component of MMIS.

All three categories of control are based on the same conceptual framework, which, in process terms, includes the functions of deterrence, monitoring, screening and intervention. The process itself can be understood through a successive pairing of these functions. Thus, the purpose of monitoring a program's administrative performance, quality or utilization patterns is in part based on the notion that "behavior observed is behavior changed." The second function of program monitoring is designed to provide the information necessary for screening or exceptions reporting. If the system is computerized, the screens will typically take the form of percentile rankings of prevailing practices or predetermined levels of acceptable behavior. Any "unusual" patterns or outcomes (e.g., cost overruns, poor quality or apparent overutilization and, conversely, cases demonstrating unusually favorable outcomes) are edited out for manual review. This review, in turn, is designed to provide greater scrutiny for the exceptional cases and thereby target the intervention strategy.

The intervention strategy is perhaps the most important part of the entire post-treatment review process, for rarely, if ever, will the deterrent effect alone provide the necessary impetus for improvement in the system under review. To be effective, the review process must be tied to an appropriate agency response: either positive incentives to induce a desired change or regulatory sanctions if the change does not come about voluntarily. The form of intervention will depend on the type of system under review and on the importance of the particular case. For example, in claims review, intervention can take the form of warnings, cancellation of participation status, retroactive denial of payment or recovery proceedings and, ultimately, criminal prosecution in the event that providers are overcharging or are guilty of other fraudulent practices. In the case of quality control, the intervention might incorporate peer pressure and educational programs designed to promote adoption of new diagnostic techniques or treatment procedures, as well as the threat of suspended hospital staff privileges for unqualified practitioners. Intervention strategies linked to utilization review could conceivably employ some of the same incentives and sanctions. More typically, these strategies would entail continued monitoring and, perhaps, the imposition of such direct utilization controls as prior authorization on identified overutilizers or overproviders.

Each of the three generic categories of post-treatment review (and attendant intervention strategies) is of obvious importance to the realization of Medicaid cost containment objectives. The results of past failures to implement effective fiscal and management review are painfully clear in terms of haphazard administration, program abuse and fraud, and excessive outlays. From a prospective standpoint, effective management review is essential to the determination of which of the various utilization controls is most cost-effective under what circumstances. Similarly, the influence of quality assurance programs on Medicaid expenditure levels cannot be ignored once objectives for cost containment are specified. However, of the three categories, utilization review is the most directly related to the substance of the present chapter.

1. Institutional Utilization Review

The requirement that medical facilities participating under Medicare and Medicaid institute utilization review (UR) committees originated with the 1965 amendments to the Social Security Act. Such committees are still required, but their functions are not largely subsumed or supervised by local PSROs. Where authority is shared by UR committees and PSROs, the purpose is to facilitate full transition to the PSRO review system. As has already been noted, the PSRO system emphasizes concurrent certification and length-of-stay review rather than post-treatment analysis.[126] Consequently, the importance once placed on post-treatment utilization review by hospital UR committees is of historical interest alone.

In point of fact, such post-treatment review was never required under federal regulations. Hospitals became involved in this type of review simply because they were unwilling or incapable of screening treatments in any other way. The actual review process varied widely from institution to institution. Some hospitals employed computer profiling to compare their own utilization patterns with PAS-LOS norms; others relied on the personal standards of appropriateness developed by UR committee members. A few institutions routinely checked all Medicare and Medicaid discharge records; most relied on unsophisticated (and typically biased) sampling methods for selecting cases for review. In some cases, the review function was handled largely by medical records librarians, in other cases solely by staff physicians. Rarely did reviewers receive special training, nor (in the case of physicians) were they paid for their review activities. Likewise, few institutions developed effective intervention strategies to cope with unusual utilization patterns once they were discovered. The outcome of the entire exercise was a confusing array of institution-specific techniques, a vast but generally worthless literature on hospital UR "success" stories, and little or no net impact on the provision of medically inappropriate or unnecessary hospital services.[127]

Different results should never have been expected. The development and implementation of institutional UR regulations violated virtually every requirement for effective utilization control: (1) objective and operational criteria for defining misutilization were not specified, (2) the choice of technique for monitoring and screening utilization patterns was left up to the individual hospital, (3) intervention strategies were undefined, (4) targeting was neither appropriate nor uniquely aimed toward cost containment, (5) enforcement was impossible given the vagueness of the regulations, and (6) effective and efficient program administration proved to be a joke. In sum, the experience with post-treatment utilization review by individual institutions provides an excellent and valuable lesson in how not to administer utilization control.

2. PSRO Profile Analysis

Although not its primary function, retrospective review of the medical necessity, quality and appropriateness of hospital care does fall under the purview of PSRO authority. The PSRO is provided with two basic mechanisms for fulfilling this function: claims review and profile analysis. In both cases, the status of current federal regulations is fluid and, in certain instances, contradictory. Regarding claims review, the PSRO *Manual* states that:

> Retrospective review of individual hospital claims is not an initially required PSRO review mechanism. It will be used only when required forms of review have not been implemented or, where implemented, have not been performed effectively.[128]

As of mid-1976, DHEW had yet to issue guidelines related either to the appropriate timing of claims review or to its relationship with other PSRO review mechanisms.[129] As a consequence, PSROs are now, in effect, free to bypass claims review altogether; given the stigma attached to retrospective denial of payment, it can be expected that most will choose to do so.

Profile analysis is another matter. As a mandated review function, profile analysis may be defined as:

> A presentation of selected information which identifies patterns of health care delivery during a defined period of time. Profile analyses will usually take the form of statistical reports produced by summarization of the data available....Profiles will allow comparisons of the patterns of care of different providers (institutions and practitioners), comparisons of current patterns with those from a previous period for the same provider, and the identification of exceptional patterns.[130]

This type of review has several objectives, including: (1) monitoring the effectiveness of concurrent review activities to see whether changes are occurring or should occur, (2) assisting in the determination of medical care evaluation study priorities, and (3) identifying area norms and providing the basis for further review of cases falling outside these norms.[131]

However, profile analyses are required only "when the capacity exists to develop them in the respective PSRO areas."[132] "Capacity" is a rather vague term, particularly in light of the fact that PSROs are not encouraged to develop their own data processing systems. An alternative, of course, is to rely on the data systems of existing carriers and intermediaries. But, as the Senate Finance Committee has pointed, out, the carriers are under no obligation to comply with PSRO requests:

> The committee would stress that the approach recommended does not envisage Blue Cross or Blue Shield or other insurance organizations or hospital or medical association review committees assuming the review responsibilities for the professional standards review organizations. Where Blue Cross or Blue Shield or other insurers or agencies have existing computer capacity capable of producing the necessary patient, practitioner, and provider profiles in accordance with the parameters and other

> requirements of the PSRO, on an ongoing expeditious and economical basis, it would certainly be appropriate to employ that capacity as a basic tool for the professional standards review organizations; but that mechanism would be employed essentially to feed computer printouts to the review organization which would be responsible for their evaluation.[133]

In those circumstances where carriers are willing to provide patient and provider profiles to local PSROs, there are the inevitable problems associated with data format, timeliness, validity and accuracy. As most state Medicaid agencies know from experience, even a formal contract between carrier and data user does not guarantee a smooth flow of information.

As a final alternative, the PSRO could rely on the state's own S/UR profiles, at least for Medicaid patients and providers. But as of April 1976, only eight states (New Hampshire, Michigan, Minnesota, Ohio, Arkansas, New Mexico, Montana and Utah) had MMIS certified systems in place;[134] this option is clearly some distance away. In the meantime, most PSROs will either have to forego profile analysis or conduct it on an ad hoc basis with simple manual tabulations derived from concurrent review reports for a few selected providers and/or patients.

While the most important impediment to effective profile review is uncertainty of data sources, there is the equally significant question of what the PSRO is to do with the data when it eventually becomes available. At present, this issue is problemmatical since federal guidelines are still being formulated. One can only hope that the mistakes evident in earlier attempts to implement utilization review procedures are not repeated.

3. Surveillance and Utilization Review Systems

At some point in time, PSRO profile analysis will undoubtedly be formally tied in with state Surveillance and Utilization Review systems; for this reason, the S/UR approach to utilization control deserves special attention.

Objectives, Standards and Data Requirements

Surveillance and Utilization Review is one of six subsystems of the Medicaid Management Information System currently being promoted by the DHEW Social and Rehabilitation Service.[135] In one sense, the S/UR system is like any of hundreds of other claims review mechanisms employed by private insurers, medical foundation plans and various governmental vendor payment programs. The basic data inputs on service delivery, utilization and expenditures all come from invoice processing. But whereas most claims review mechanisms are designed primarily to check the validity and accuracy of provider invoices and recipient eligibility status, these two functions are handled elsewhere in the MMIS system. S/UR is designed, rather, to serve the broader objective of reducing misutilization and program abuse through (1) monitoring service delivery and utilization patterns, (2) establishing uniform and, where possible, nonnormative standards of care, (3) screening or exceptions reporting based upon these standards, (4) targeting direct intervention mechanisms when standards are not met, and (5) providing, in the process, a deterrence to future misutilization and abuse.

Over time, the monitoring function of S/UR is intended to generate five interrelated sets of statistical reports.[136] The first presents an historical profile for each Medicaid participant, including utilization experience for eligible program recipients and service delivery rates for program providers. The second is a summary report which profiles service usage rates among recipients grouped according to similar demographic, medical and utilization characteristics. The third is a claim detail report which provides specific information on diagnosis codes, procedures codes and dates of service. These data are used in turn to develop the treatment summary analysis report which profiles service delivery patterns for specific diagnoses. The system is also designed to produce a fifth category of special reports which will test the value of alternative profile formats and handle special needs not met by the standard reports.

The profile reports lay the groundwork for standard setting under S/UR. They make it possible to establish ranges for "usual" patterns for the particular recipient class, provider class or diagnosis (statistical norms), but they may also be supplemented by normative medical opinion (medical standards). In either case, the process has two distinct advantages over standard setting under present forms of direct utilization control, specifically, over concurrent review. First, the baseline norms are generated within the system, avoiding the validation problem of translating standards from one population group to another. Second, the system can establish norms both on a micro, treatment-specific basis and on a macro, class-specific basis. Treatment specific norms are, of course, employed under concurrent review, but at present they apply only to hospital length of stay; given the rather slow pace of PSRO development, it will be some time before norms addressing other aspects of treatment are produced. The S/UR system can generate norms for LOS, for other inpatient procedures (e.g., number and type of laboratory tests prescribed by diagnosis) and for ambulatory services (e.g., frequency of office revisit by diagnosis). Moreover, when it comes down to choosing which treatment-specific standards will be applied under a direct utilization control mechanism, the class profiles can be used to generate impact estimates in terms of cost and resource use for a variety of alternatives. This capacity provides an essential prerequisite to informed decisions regarding quality/cost trade-offs in any utilization control program.

Once the standards are determined, a system of machine edits is employed to flag exceptional cases for manual review. The S/UR exceptions reporting mechanism is equally suited to flagging apparent cases of underutilization as it is to citing overutilization and program abuse (e.g., tonsillectomies with an average inpatient stay of 2 or 8 days, when the age-specific standard is 4 days). The actual range of exceptions reported over and below a given standard will ultimately depend on the cost-effectiveness of the manual review procedure. If it turns out, for example, that a review of all cases within one standard deviation of the class or treatment standard indicates a very low level of unjustified utilization, then the band of acceptable variation can be widened.

Given the characteristics of exceptions reporting, the S/UR system will only identify major anomalies in treatment procedures or in delivery and utilization patterns. As it cannot determine causation, an investigations unit is necessary for this function. The unit might employ a wide variety of investigatory techniques, depending on the nature of the case. A "typical" instance of apparent overprescription by a participating Medicaid physician might, for example, entail an in-house review of that physician's billing history, a verification that recipients actually received the prescriptions, a field audit of the doctor's medical records and, perhaps, an audit of pharmacies which filled the prescriptions. The information obtained through the investigation would then be used to determine the state's course of action. If the physician was found to be practicing bad medicine, the case might be referred to the local medical society or, if serious enough, to the state medical licensure board. Alternatively, the physician could be subject to a retroactive denial of payment, targeted for ongoing review or prior authorization, or simply dropped from the program. If the physician were guilty of receiving fraudulent kick-backs from local pharmacies, the state could be expected to seek recovery of funds and criminal prosecution.

The most obvious function of investigation and intervention is to discover the causal factors underlying misutilization and abuse and to correct the deficiencies once uncovered. Remedial action, however, is not the only way to effect change. All forms of utilization control have some deterrent influence on provider and/or recipient behavior, but since most are targeted to a few specific services, that influence will not extend far and may, in fact, simply shift the problem elsewhere. The S/UR system, by profiling the entire range of service delivery and utilization, thus has a greater potential deterrent effect.

Current Regulatory Status

All payors under Titles XIII and XIX (carriers, intermediaries and state Medicaid agencies acting as their own fiscal agents) are required to perform certain minimal postpayment surveillance and profiling activities as a routine function of claims review. For the most part, these activities take the form of monthly or quarterly reports which provide little more than aggregate service-to-patient ratios. Only in the case of physicians are payers required to profile individual patterns of practice or service delivery.[137] Payers are encouraged to go beyond minimum reporting requirements, but given wide differences in EDP capacity, it is simply unrealistic to require the kind of detail necessary to perform most S/UR functions.

In part for this reason (and in part to generally augment state agency capabilities in the area of mechanized claims processing and information retrieval systems), Title XIX has, since 1974, provided increased federal financial participation (FFP) for states implementing and operating MMIS systems. The FFP for the design, development and installation of new mechanized systems (or improvements to existing systems) is 90 percent of program costs. A 75 percent FFP is provided for operating expenses of certified MMIS systems.

Current Users and Adoption Implications

With such generous federal financing and the potential for significant program cost savings when implemented, it is amazing that so few states have demonstrated an interest in MMIS. As noted previously, only eight states have certified systems in place. An additional ten states were developing MMIS systems under approved state implementation plans as of mid-1976 (thereby eligible for 90 percent FFP), but a full 17 states were either inactive as of that time or had shown no interest in participating.[138] A few states have developed their own systems of utilization profiling, but none are as extensive as S/UR. In New York, for example, the NYSHUR (New York Hospital Utilization Review) system provides experience-generated norms by diagnosis, age and sex for institutional utilization patterns.[139]

The first state to implement MMIS (and hence S/UR) was New Mexico in May 1975. Five of the remaining seven states implemented their systems in 1976. There is, in short, no experiential basis for evaluating the effectiveness of S/UR in curtailing misutilization and abuse under Medicaid. Why, then, should it be adopted? There are five major reasons. First, S/UR provides the state with a holistic picture of the service delivery and utilization patterns under its Medicaid program. It must be remembered that cost containment through utilization control is a secondary objective, to be met subject to attainment of the program's primary goals. The S/UR system lays the groundwork for an assessment of whether these goals have been reached.

Second, the system offers a mechanism for rationally targeting utilization control by service, provider and recipient categories. Up to this point, states have haphazardly implemented a raft of different controls based on little solid evidence of where control is most needed. Thus S/UR, in conjunction with its profiling functions, provides a relatively inexpensive method of establishing uniform norms and standards of service delivery and utilization. The lack of consistent standards, a major weakness of pretreatment and concurrent review, could be rectified if such controls were incorporated as intervention strategies under S/UR.

Fourth, the Surveillance and Utilization review system provides a basis for utilization control effectiveness studies. The need for in-depth empirical evaluation of both the type and level of control activity should be obvious, but such evaluation cannot be done without adequate baseline profile data. Not only will the system produce such a data base over time; more importantly, the data elements can be readily analyzed in relation to control norms and standards. This consideration leads to the final rationale for S/UR adoption. As stressed repeatedly, the development of an effective utilization control program is conditioned upon concise objectives for cost containment, subject to maximum quality constraints. In these terms, S/UR sets the stage for a solution to the cost/quality conundrum.

H. Conclusions

The need for utilization control under Medicaid is based on two assumptions: first, that the extent of misutilization in current systems of medical care delivery represents a serious drain on scarce resources, and second, that regulatory rather than market reform is the appropriate public response. This chapter has summarized (from the perspective of cost containment strategy) the considerable circumstantial evidence of medical misutilization. Sidestepped, for the most part, was an equally extensive literature supporting the argument that a greater, as opposed to a lesser, utilization of medical services is in the public interest. The omission was purposeful. Given the considerable ambiguity in measuring social values for health care expenditures, any determination of "optimal" levels of medical spending is fundamentally political rather than economic in character. Anyone in the least bit familiar with the direction of state Medicaid policy in the last few years knows that the political process is likely to force a further retrenchment in state program efforts.[140] Under these circumstances, a call for expansion of coverage or quality of care will be dismissed out of hand in most states. More importantly, such a call may prove counterproductive in that it shifts emphasis away from activities (such as utilization control) which aim to maintain current levels of program effectiveness within the framework of a reduced rate of increase in program outlays.

This chapter has also omitted any discussion of the recurrent debate among economists and political theorists over the relative effectiveness of regulation compared to increased competition among private health providers.[141] Again, the reasons are pragmatic. While most experts in this area can agree that the health sector is "overregulated" in the aggregate, there is considerably less agreement when it comes to specific aspects of the delivery system. On the one hand, it is relatively easy to construct a case for regulation on such traditional grounds as natural monopoly (e.g., hospital services in medium- and small-sized towns), the public interest doctrine and consumer ignorance. But, on the other hand, it can be argued that the circumstances which support regulation are not inherent in the system but, rather, the product of past regulatory activities (e.g., professional and institutional licensure and prohibitions on advertising). However, the point is that no publicly elected or politically appointed body is going to oversee the wholesale dismantling of current regulatory machinery, regardless of the merits of the case. Not only are provider interest groups and consumer lobbies powerful enough to preclude such action; in a very real sense, the effects of deregulation (and hence the wisdom of the action) are essentially unknown. Even if one could prove the ultimate value of deregulation, the transitional impact might well be so disruptive as to negate all but the most carefully staged long-range planning effort.

In the meantime, it is necessary to design intervention strategies which do not exceed current political and economic constraints. It is within this context that direct methods of utilization control can be advocated as means of improving the cost-effectiveness of government Medicaid expenditures. There is no question that utilization control has this potential; the issue, rather, is how best to maximize it. Various mechanisms have been discussed throughout this chapter but, because of a general lack of baseline data and a reliance on inadequate methodological procedures, their true potential is difficult to estimate.

Therefore, before a state sets a course involving one or more specific control mechanisms, top priority should be directed toward the information component of utilization control in general. Two interrelated efforts are required in this regard. First, the state should establish a means of monitoring current utilization patterns in order to formulate realistic future standards and norms. For reasons outlined in the last section, the implementation of MMIS and its S/UR subsystem represents a useful initial step in this direction. But a functioning information retrieval system is not enough. While the process of setting utilization standards and norms is a political activity,[142] public policy should be based on an accurate assessment of the monetary trade-offs between quality of care and program expenditures. Since the returns to "quality" medical care are nearly impossible to measure in dollar terms, the only realistic alternative is to set an arbitrary maximum quality constraint and then minimize program expenditures subject to that constraint.

This alternative presupposes, of course, that the savings realizable from alternative utilization control mechanisms can be measured. The development of an accurate and uniform methodology for this purpose is the second major effort required before a state can realistically choose which controls will be the most effective. While such a methodology has yet to be fully fleshed out, several important aspects have been discussed in the body of this chapter, including: (1) the necessity for measuring recipient and provider response to each control mechanism in terms of alternative service delivery options, (2) the separation of transitory and/or spurious effects from real program effects, (3) the impact of each control mechanism over time, (4) the generalizability of program impacts across different recipient, provider and/or geographical groups, (5) the need to incorporate the administrative costs of each option in the measurement of cost-effectiveness, and (6) the need to separate short-run and long-run savings and to apply appropriate marginal cost estimating procedures in each case.

Unfortunately, what a state should and can do in designing effective utilization controls are two separate issues. One can argue that sophisticated information systems and sound estimating methodologies are essential if the state is to avoid the consequences of ad hoc, trial-and-error decision making. However, state government must still contend with the political pressure to attack the problems of Medicaid misutilization and program abuse before such systems can possibly be implemented. The immediate options are limited. Federal regulations specify that state Medicaid agencies must comply with Title XIX requirements for institutional UR and PSRO review activities, both of which place a heavy emphasis on concurrent review. From the perspective of cost containment, it is indeed regrettable that this form of utilization control may ultimately result in increased rather than reduced program expenditures. Counter-balancing the potential expansionary impact of PSRO activities will require the combined application of pre-delivery controls and ex-post targeting.

As mentioned repeatedly, the cost-effectiveness of any control mechanism can be increased through selective targeting of intervention strategies. Up to the present, the most common form of state intervention has been through the power of retroactive denial of payments. Because this form of intervention is precluded by PSRO concurrent review activities, the states are left with the option of preventing misutilization before it occurs. Maximizing the return from predelivery controls, however, requires effective monitoring and specific targeting. Clearly, states with operating S/UR systems have a distinct advantage when it comes to identifying providers, recipients and service categories for direct control. But even states without such systems can apply the following rules of thumb.

First, targeting should be based on the expected dollar return from controlling utilization and not on the probability of misutilization per se. It generally does not pay to regulate the provision of low-cost services, even when the probability of misutilization may be relatively high. In particular, the expected returns will be highest for hospital services and nursing home care (especially in states with high bed capacities), for specialist rather than generalist care, and for recipients with high rates of utilization as opposed to recipients with average or low utilization patterns.

Second, the cost of administering the control should be low relative to the expected savings. Of the two basic types of predelivery controls, prior authorization is clearly preferable to such prior review mechanisms as second consultations for elective surgery. Not only does a mandatory authorization requirement eliminate the problems of noncompliance and thereby increase the expected dollar return; the requirement is also less expensive to administer because of centralized staffing and the application of uniform standards. However, it should be noted that the per unit cost of administration for any control will rise as the targeting becomes increasingly case- or group-specific. The level of cost-effective specificity cannot be determined by any simple rule of thumb and will thus depend on experience.

Finally, any system of targeting pre-delivery controls should be flexible enough to adjust to changing circumstances. The system should be managed so that direct control mechanisms are replaced (perhaps by utilization monitoring) wherever the specific targeting ceases to be cost-effective, as might be expected when provider and/or recipient behavior adjusts to the control mechanism. Likewise, the system should be flexible enough to impose controls when such targeting would indicate a positive net return.

NOTES TO CHAPTER I

1. Some of the literature review presented in this section was adapted from a memorandum prepared by Bruce Spitz, now a colleague at The Urban Institute, while he was a health policy analyst with the State of Michigan. This section has benefited from those comments as well as from subsequent discussions.

2. Byron Straight, "Reducing the Incidence of Office and Home Visits in a Medical Service Plan by Use of Coinsurance Charges," Proceedings of a Conference of Actuaries in Public Practice, Vol. 11, pp. 73-79.

3. R. G. Beck, "The Effects of Copayment on the Poor," Journal of Human Resources, Vol. 9, No. 1 (Winter 1974), pp. 129-141.

4. Saskatchewan Insurance Commission, Annual Report, 1970, Table 3.

5. Charles E. Phelps and Joseph P. Newhouse, "The Effects of Coinsurance on Demand for Physician Services," Rand Corporation, Report R-976-OEO (Santa Monica, California).

6. Ann Scitovsky and Nelda M. Snyder, "Effects of Coinsurance on Physician Services," Social Security Bulletin, Vol. 35 (June 1972), pp. 3-19.

7. Milton I. Roemer, M.D., Carl E. Hopkins, Ph.D., Lockwood Carr, B.S., and Foline Gartside, M.S., "Copayments for Ambulatory Care: Penny-Wise and Pound-Foolish," Medical Care, Vol. XIII, No. 6 (June 1975).

8. Ibid.

9. Carl E. Hopkins, Milton I. Roemer, et al., "Cost Sharing and Prior Authorization Effects on Medical Services in California: The Providers' Reactions," Medical Care, Vol. XIII, No. 8 (August 1975).

10. Collection problems of physicians would be eliminated if copayments were collected by the states. For example, if the states collected cost sharing liabilities by reducing welfare payments for cash assistance recipients and by billing medically needy recipients, providers would be free from those burdens. However, such an approach suffers in that the costs of adjusting welfare payments and the amount of uncollected patient liabilities may both be substantial. Moreover, individuals, even in the long run, are unlikely to be responsive to copayments collected retroactively.

11. John Holahan and Bruce Spitz, "Physician Reimbursement," in Altering Medicaid Reimbursement Arrangements (Washington, D.C.: The Urban Institute, 1977).

12. Robert D. Evans, "Supplier-Induced Demand: Some Empirical Evidence and Implications," in Mark Perlman (ed.), The Economics of Health and Medical Care (New York: Halstead Press, 1974).

13. Charles P. Hall, "Impact of Cost-Sharing on Consumer Use of Health Services," in Selma J. Mushkin, Consumer Incentives for Health Care (New York: Melbach Memorial Fund, 1974).

14. For reviews of the literature on cost sharing and hospital utilization, see Sylvester Berki, Hospital Economics (Lexington, Massachusetts: Lexington Books, 1972), Chapter 6, and Robert Kaplan and Lester E. Lave, "Patient Incentives and Hospital Insurance," in Health Services Research (Winter 1971), pp. 288-300.

15. Bruce Spitz, "Hospital Reimbursement," in Altering Medicaid Reimbursing Arrangements (Washington, D.C.: The Urban Institute, 1977).

16. John Rafferty, "Patterns of Hospital Use: An Analysis of Short-Run Variations," Journal of Political Economy, Vol. 79, pp. 154-65.

17. Information supplied by the Medical Services Administration.

18. Philip Lee and Milton Silverman, Pills, Profits and Politics (Berkeley: University of California Press, 1974).

19. Ibid., p. 305.

20. John Holahan, Physician Supply, Peer Review and Use of Health Services in Medicaid (Washington, D.C.: The Urban Institute, 1976).

21. Hearing before the Subcommittee on Monopoly of the Select Committee on Small Business, Part E, December 5-14, 1972.

22. Charles E. Phelps and Joseph P. Newhouse

NOTES TO CHAPTER II

1. See, for example, Larry Frederick, "How Much Unnecessary Surgery?" Medical World News (May 3, 1976), pp. 50-66; Bruce Stuart and Ronald Stockton, "Control Over the Utilization of Medical Services," Milbank Memorial Fund Quarterly: Health and Society, Vol. 51, No. 3 (Summer 1973), pp. 341-394; and Carl Slater and Alice Bryant, "Regulation of Professional Performance" in Effects of Regulation in the Health Sector (Denver: Policy Center, Inc., 1974).

2. E.G. McCarthy and G.W. Widmer, "Effects of Screening by Consultants on Recommended Effective Surgical Procedures," New England Journal of Medicine (December 19, 1974), pp. 1331-1335; and Statement by Eugene McCarthy in Subcommittee on Oversight and Investigations of the House Committee on Interstate and Foreign Commerce, 94th Congress, 1st Session, Hearings on Getting Ready for National Health Insurance: Unnecessary Surgery (1975), p. 133.

3. Subcommittee on Oversight and Investigations, Hearings on Getting Ready, p. 133.

4. See Ingelfinger, "Misuse of Information Printed in the Journal," (editorial) New England Journal of Medicine, Vol. 294 (March 18, 1976), p. 667; New York Times, February 6, 1976, p. 1, col. 1; and letter to Hon. John E. Moss, September 15, 1975, from Theodore Cooper, Assistant Secretary for Health, published in Subcommittee on Oversight and Investigations of the House Committee on Interstate and Foreign Commerce, 94th Congress, 2nd Session, Report on Cost and Quality of Health Care: Unnecessary Surgery (Subcommittee Print, 1976), pp. 45-46.

5. Subcommittee on Oversight and Investigations, Report on Cost and Quality, pp. 30-31.

6. Ibid, pp. 31-32.

7. Jane Brody, "Cornell Study Calls 11 Percent of Surgery Unnecessary," New York Times, May 3, 1975, p. 58c.

8. Subcommittee on Oversight and Investigations, Report on Cost and Quality, p. 30.

9. Wennberg, Subcommittee on Oversight and Investigations, Hearings on Getting Ready, p. 76.

10. Marian Gornick, "Medicare Patients: Regional Differences in Length of Hospital Stays 1969-1971," Social Security Bulletin, Vol. 38, No. 7 (July 1975), p. 16.

11. See, for example, John Holahan, "Physician Supply, Peer Review and Use of Health Services in Medicaid," Urban Institute Paper 976-6 (Washington, D.C.: The Urban Institute, February 1976); J.P. Bunker, "Surgical Manpower, A Comparison of Operations and Surgeons in the United States, England, and Wales," New England Journal of Medicine, Vol. 285 (January 1970), pp. 135-144; and E. Vayda and G.D. Anderson, "Comparison of Provincial Surgical Rates in 1968," The Canadian Journal of Surgery, Vol. 18 (January 1975), pp. 18-26.

12. U. S. Department of Health, Education and Welfare, Public Health Service, Federal Employees Health Benefit Program Utilization Study (Washington, D.C.: GPO, 1975).

13. See, for example, the studies reviewed in Milton Roemer and William Shonick, "HMO Performance: The Recent Evidence," The Milbank Memorial Fund Quarterly: Health and Society, Vol. 51, No. 3 (Summer 1973), pp. 271-317; E.F.X. Hughes et al., "Utilization of Surgical Manpower in a Prepaid Group Practice," Surgery, Vol. 77 (March 1975), pp. 371-383; C.R. Gaus et al., "Contrasts in HMO and Fee-for-Service Performance," SSA Office of Research and Statistics, December 30, 1975 (mimeograph); and Council on Wage and Price Stability, Employee Health Care Benefits, preliminary draft printed in the Federal Register, September 17, 1976, Part III, pp. 40312-40315.

14. See references under note 1.

15. M.O. Keplar, "The Physician and Drugs: Professional Misuse and Personal Abuse," Man and Medicine, Vol. 1 (Winter 1976); and Boyce Rensberger, "Bad Prescriptions Kill Thousands a Year," New York Times, January 28, 1976, p. 1c.

16. Edward Wagner et al., "Influence of Training and Experience on Selecting Criteria to Evaluate Medical Care," New England Journal of Medicine, Vol. 294, No. 16 (April 15, 1976), pp. 871-876; and Laurence Altman, "Study Finds that Much Use of Antibiotics May Not be Justified," New York Times, April 8, 1976, p. 22c.

17. Council on Wage and Price Stability, Employee Health Care Benefits, pp. 40303-40305.

18. Odin Anderson and Paul Sheatsley, "Hospital Use: A Survey of Patient and Physician Decisions," Research Series, No. 24, Center for Health Administration (Chicago: University of Chicago, 1967).

19. Subcommittee on Oversight and Investigations, Hearings on Getting Ready, p. 185.

20. Felicia Holton, "What Future Now for Surgical Centers?" Prism (March 1975), p. 52.

21. See, for example, the studies cited in Daniel Bernstein and Richard Egdahl (eds.), "Peer-Review Components of the Health-Care System," Boston University Medical Center, Perspectives on Health Policy, Vol. 1 (November 1974).

22. James Zimmer, "Length of Stay and Hospital Misutilization," Medical Care, Vol. 12, No. 5 (May 1974), pp. 453-462.

23. Paul Gertman, "Hospital Length of Stay," in Bernstein and Egdahl, Perspectives on Health Policy, p. 58.

24. Clifton Gaus, Findings presented before the National PSRO Advisory Council, March 9, 1976, reported in "Hospital Costs, Stays Vary Widely Across U.S.," American Medical News, Vol. 19, No. 1 (March 22, 1976).

25. Gornick, "Medicare Patients," p. 22.

26. John Holahan and Christian Ehemann, "Public Policy and Utilization of Long Term Institutional Care," Urban Institute Paper 975-08 (Washington, D.C.: The Urban Institute, 1976).

27. For example, the most common source of information detailing length of stay differences among hospitals is found in the CPHA Professional and Hospital Activities (PAS) reports. While the number of hospital days reported by PAS tends to be fairly reliable, there are indications that upwards of 40 percent of principal diagnoses records on hospital case abstracts (which form the basis for the PAS data system) may be incorrect for particular institutions, and that the average national error rate for principal diagnoses runs between 12 and 15 percent. Naturally, errors of this type will bias any cross-sectional estimates of length of stay differences by diagnostic category. See Gertman, "Hospital Length of Stay," p. 55.

28. Clark Havighurst and James Blumstein, "Coping with Quality/Cost Trade-Offs in Medical Care: The Role of PSROs," Northwestern University Law Review, Vol. 70, No. 1 (March-April 1975), p. 17.

29. Ibid, pp. 38-60.

30. See, for example, "Nursing Homes Now Face Dozen Inquiries on Abuses," New York Times, January 12, 1975, p. 41; Edith Asbury, "Hollander Admits Fraud on Medicaid in Nursing Homes," New York Times, February 3, 1976, p. 1; John Hess, "The Eternal Nursing-Home Inquiries," New York Times, May 30, 1976, p. 6c; and Statement by Senator Domenici, "Medicare/Medicaid Abuses," in Congressional Record, Vol. 122 (February 18, 1976), p. 122.

31. See, for example, Philip O'Connor, "Medicaid Fraud - 'An Octopus'," Chicago Daily News, June 3, 1976; Dolores Katz, "Card is Key to Assembly-Line Care," Detroit Free Press, July 22, 1975, p. 1; Saul Friedman, "State Fraud

Lists Medicaid Abuses," Detroit Free Press, September 27, 1975, p. 2a; and Robert Longstaff, "State Crackdown Costs Doctor's Medicaid Take," in Ann Arbor News, February 2, 1976.

32. See, in general, U.S. General Accounting Office, History of the Rising Costs of the Medicare and Medicaid Programs and Attempts to Control These Costs 1966-1975, Report to the Human Resources Task Force, House Committee on the Budget (Washington, D.C.: GPO, February 11, 1976); and "State Plan Requirements" and "New Developments" sections in Medicare and Medicaid Guide, Commerce Clearing House, Inc. (updated biweekly).

33. D.R. Bailey and D.C. Riedel, "Recertification and Length of Stay," Blue Cross Reports, Vol. 4 (July 1968).

34. E. Brian, "Foundation for Medical Care Control of Hospital Utilization: CHAP--a PSRO Prototype," New England Journal of Medicine, Vol. 288 (April 26, 1973).

35. Rona Sayetta, "Critique of an Earlier Study of the Sacramento Medical Care Foundation's Certified Hospital Admission Program (CHAP)," Medical Care, Vol. 14, No. 1 (January 1976), pp. 80-90.

36. Subcommittee on Oversight and Investigations, Cost and Quality, p. 32.

37. See discussion in Council on Wage and Price Stability, Employee Health Care Benefits, pp. 40300-40303.

38. Ibid, p. 31.

39. Donna Barbaro, Larry Shuman, and Robert Swinkola, "An Evaluation of Various Presurgical Testing Procedures," Research Series, No. 16, Blue Cross of Western Pennsylvania, January 1976, p. 6.

40. Indirect evidence suggests that inefficiencies in hospital administration may be a more important cause of misutilization than inappropriate diagnostic admissions and pre-operative overstays. See Barbaro et al., "Evaluation of Testing Procedures," pp. 39-48.

41. 42 U.S.C. § 1396a (a)(19)(1970).

42. Medicare and Medicaid Guide, p. 5245.

43. Procedure Manual, "History and Summary of the PSRO Legislation," Western Massachusetts PSRO Inc., July 1976, p. 3.

44. 42 U.S.C. § 1320c-4(a)(2)(Supp. II, 1972).

45. U.S. Department of Health, Education and Welfare, Office of Professional Standards Review, PSRO Program Manual (March 15, 1974) § 705.13.

46. Barbaro et al., "Evaluation of Testing Procedures," pp. 3-6.

47. For example, the "Program for Consultation on Elective Surgery," PROCESS, was initiated by Michigan Blue Cross in September 1975; the "Program for Elective Surgical Second Opinion" or PRESSO in New York and SCP "Surgical Consultation Program" in Pennsylvania both commenced operation on January 1, 1976. See Council on Wage and Price Stability, Employee Health Care Benefits, pp. 40319-40320.

48. Ibid, p. 40319.

49. Ibid, pp. 40319-40320.

50. Ibid, p. 40303.

51. Barbaro et al., "Evaluation of Testing Procedures," pp. 4-5.

52. Ibid, pp. 6-7.

53. Brody, "Cornell Study," p. 58.

54. Council on Wage and Price Stability, Employee Health Care Benefits, p. 40303.

55. Ibid, pp. 40301-40302.

56. Barbaro et al., "Evaluation of Testing Procedures," p. 17.

57. Ibid, pp. 39-48. See also Karin Dumbaugh and Duncan Newhauser, "The Effect of Pre-Admission Testing on Lengths of Stay," in Stephen Shortell and Montague Brown (eds.), Organizational Research in Hospitals (Chicago: Blue Cross Association, 1976).

58. Council on Wage and Price Stability, Employee Health Care Benefits, p. 40302.

59. Ibid, p. 40301.

60. U.S. Senate, Finance Committee Staff Report, Medicare and Medicaid-Problems, Issues, and Alternatives (Washington, D.C.: GPO, February 1970), p. 128.

61. U.S. Department of Health, Education and Welfare, SRS, "Utilization Review: Grants to States for Medical Assistance," Federal Register, Tuesday, March 30, 1976, Part II, p. 13452.

62. U.S. Department of Health, Education and Welfare, SRS, MSA, Medical Assistance Manual, §§ 5-30-10(A-C)-5-30-20(A-6), referenced in Medicare and Medical Guide, p. 6414.

63. Ibid, p. 6415.

64. Ibid, pp. 6415-2.

65. Council on Wage and Price Stability, Employee Health Care Benefits, p. 40304.

66. Ibid, p. 40320.

67. State of California, Department of Health, Field Services Section, Prior Authorization Study Project Final Report, March 24, 1975, p. A.1.

68. Ibid, pp. 26, 28.

69. Wisconsin Department of Health and Social Services, Wisconsin's Medical Assistance Program: Rising Costs and Cost Containment, January 1, 1976, p. 17.

70. Ibid, pp. 23-24.

71. Pennsylvania Department of Public Welfare, Medical Assistance Manual, Section 9700, "Prior Approval of Selected Services and Payments," July 1, 1975.

72. Katharine Bauer, "Containing Costs of Health Services Through Incentive Reimbursement: Some Approaches by Third Party Payers," Cases in Health Services, Series No. 4. (Cambridge, Massachusetts: Harvard Center for Community Health and Medical Care, December 1973), pp. 214-216.

73. Bruce Flashner et al., "The Hospital Admissions and Surveillance Program in Illinois," JAMA, Vol. 221 (September 4, 1972), p. 1156.

74. Council on Wage and Price Stability, Employee Health Care Benefits, p. 40304.

75. Ibid.

76. E. Brian, Government Control of Hospital Utilization: A California Experience (Sacramento: California Department of Health Care Services, 1971); and California Health Data Corporation, Examination of Average Stay Patterns: Hospitals in Sacramento County, January 1969-December 1970 (San Francisco: The Corporation, 1971).

77. David Kinzer, "Hospital Association Programs for Quality Assurance," in Bernstein and Egdahl, "Peer-Review Components," p. 63.

78. State of California, Prior Authorization Study, pp. 17-24.

79. Ibid, pp. 17-18.

80. Ibid, pp. 19-20. It should be noted, however, that savings to the Medi-Cal program will exceed these rates to the extent that the overhead costs associated with nonutilization are spread across the entire hospital and nursing home patient load. It can be expected, in other words, that some of the program "savings" simply reflect costs transferred to the private sector and other government programs.

81. Ibid, p. 24.

82. Ibid, pp. 21, L.5.

83. Ibid, pp. 21, I.3, I.4.

84. Ibid, p. 21 and Sections I, J, K.

85. Ibid. See discussion on pp. 1-17.

86. U.S. General Accounting Office, History of the Rising Costs, op cit.

87. All figures relate to Medi-Cal experience during 1974 with one exception; the average prescription price is the 1974 national average as reported by the Lilly Digest. See: DHEW, SSA, ORS, Prescription Drug Data Summary 1974 (Washington, D.C.: Government Printing Office, 1976), p. 37.

88. See, for example, the Pennsylvania guidelines relating to prior approval of medical equipment purchases, Pennsylvania Department of Public Welfare, section 9704.3.

89. Carl Hopkins et al., "Cost-Sharing and Prior Authorization Effects on Medicaid Services in California: Part II: The Providers' Reactions," Medical Care, Vol. 13, No. 8 (August 1976), pp. 643-647.

90. Ibid, p. 645.

91. Medicare and Medicaid Guide, p. 6415-2.

92. Ibid, p. 5227.

93. PSROs are granted authority to conduct three major review activities: (1) admission certification and continued care review, (2) medical care evaluation studies and (3) analysis of hospital, practitioner and patient care profiles (Medicare and Medicaid Guide, p. 5248). Admission certification tends to be perfunctory unless tied to prior authorization--which at present is not required. The medical care evaluation studies are supposed to provide the basis for establishing standards of medical care practice within institutions, but at present the only requirement promulgated by the PSRO is that every hospital have at least one such study underway at all times. Nothing is said about what type of study should have priority or what must be done with the study results once completed.

94. Medicare and Medicaid Guide, p. 5251.

95. Ibid, pp. 5145 and 5248.

96. Ibid, p. 5253-3.

97. Gertman, "Hospital Length of Stay," pp. 55-56.

98. Ibid, p. 59.

99. Medicare and Medicaid Guide, p. 6407.

100. Ibid, p. 5148.

101. PSRO Letter: A Report from Washington on Professional Standards Review Organizations, Vol. 4, No. 18 (September 14, 1976), p. 2.

102. Medicare and Medicaid Guide, p. 5242.

103. Western Massachusetts PSRO, Inc., Newsletter, Vol. 3, No. 4 (September 1976), p. 1.

104. For a listing of several recent plans for hospital review by employer and employee health programs, see Council on Wage and Price Stability, Employee Health Care Benefits, pp. 40319-40320.

105. See U.S. Department of Health, Education and Welfare, National Center for Health Services Research and Development, Experimental Medical Care Review Organization (EMCRO) Programs, Publication No. (MSM) 73-3017 (Washington, D.C.: GPO, March 1973).

106. Bauer, "Containing Costs," p. 232.

107. There are a number of reasons why Medi-Cal clients under CHAP might have shorter lengths of stay that are essentially unrelated to the plan's certification and concurrent review procedures: (1) the average LOS in California is some 10 percent lower than for the nation as a whole; (2) the Medi-Cal population tends to be younger and hence may use services at a lower level of intensity; (3) the CHAP data for April-August 1970 were compared with the PAS average for 1969, and LOS averages were falling during the entire period; (4) the Medi-Cal population reflects a different socioeconomic mix from the general population, which may result in divergent utilization trends; and (5) the single most frequent surgical procedure performed on CHAP Medi-Cal patients was abortion with an average length of stay of 1.9 days (at the time of the study, abortions were still illegal in other states). See Bauer, "Containing Costs," pp. 232-233.

108. Brian, "Foundation for Medical Care Control," pp. 878-882.

109. Ibid.

110. See, in particular, Sayetta, "Critique of an Earlier Study," pp. 80-90; Slater and Bryant, "Regulation," pp. 80-83; and Bauer, "Containing Costs," pp. 234-236.

111. Sayetta, "Critique of an Earlier Study," p. 87. The administrative costs per hospital admission reviewed were $10.02 and $3.89 respectively.

112. Neither estimate takes into account (1) higher hospital per diem costs for shorter average LOS, (2) increased utilization of ambulatory services in lieu of hospitalization or (3) costs associated with health loss on behalf of patients inappropriately denied services (see Sayetta, "Critique of an Earlier Study," pp. 88-89). Moreover, according to a recent survey of physician attitudes regarding the effect of concurrent review, 41 percent believed that the review mechanism would increase medically unnecessary diagnostic tests ordered primarily so that doctors could avoid the embarrassment of having their judgment overturned by a review committee (see Frederick, "PSROs: How the First Ones are Working," Medical World News (October 25, 1974), pp. 70 and 77).

113. Sayetta, "Critique of an Earlier Study," p. 87.

114. Ibid, computed from county figures shown on page 86.

115. Cedric Chao, "Cost and Quality Control in the Medicare/Medicaid Program: Concurrent Review," Harvard Civil Liberties Law Review, Vol. 11, No. 3 (Summer 1976), p. 676.

116. See references under note 110. See also Alan Nelson, "Relation Between Quality Assessment and Utilization Review in a Functioning PSRO," New England Journal of Medicine, Vol. 293, No. 13 (March 27, 1975), p. 675. Nelson relates that early experience under the OSCHUR program indicates possible reductions in average LOS of 10 percent, as opposed to the 16 percent reduction attributed to CHAP.

117. Frederick, "PSROs," p. 56.

118. U.S. House of Representatives, Ways and Means Committee, Subcommittee on Health, Hearing on Selected Issues in Medicare Program Policy; 94th Congress, 1st Session (June 12, 1975), p. 260.

119. Ibid, pp. 120-121.

120. Current regulations require that review take place within one working day of admission, but, as yet, they have not been enforced. Moreover, the "one working day" allowance means that weekend admissions will not be reviewed for two to three days.

121. See note 111.

122. Medicare and Medicaid Guide, p. 5257.

123. Ibid, p. 5258.

124. Bennett, "Professional Standards Review Organizations--Philosophy and History," Utah Law Review, Vol. 335 (1975), pp. 358-359.

125. Chao, "Cost and Quality Control," p. 680.

126. The primary exception is in the case of quality assurance (rather than utilization review as such) whereby hospitals are required to undertake medical care evaluation studies on a retrospective basis.

127. See, for example, the literature reviewed by Slater and Bryant, "Regulation," and Stuart and Stockton, "Control over the Utilization,"

128. Medicare and Medicaid Guide, p. 5266.

129. Ibid.

130. Western Massachusetts PSRO, Inc., "Objectives of Review Process," p. 2.

131. Ibid, pp. 2-3.

132. Medicare and Medicaid Guide, p. 5241.

133. Ibid, quoted on pp. 5266-5267.

134. State of Michigan, DSS, Bureau of Medical Assistance, Michigan Medicaid: MMIS, S/UR, Third Party Liability, and Medicaid Investigations (Lansing: July 23, 1976), p. 4.

135. The other components of MMIS include: (1) the Recipient Subsystem, (2) the Provider Subsystem, (3) the Claims Processing Subsystem, (4) the Reference File Subsystem and (5) the Management and Administrative Reporting Subsystem. For a description of all subsystem components, see: U.S. Department of Health, Education and Welfare, SRS, MSA, Medical Assistance Manual, Sections 7-71-40 through 7-71-60, effective June 10, 1974.

136. These profiles require more data than is typically available from provider invoices (particularly in regard to recipient demographics and provider specialty information). For this reason, S/UR is tied in with the Recipient and Provider Subsystems of MMIS.

137. Medicare and Medicaid Guide, p. 5202.

138. State of Michigan, Michigan Medicaid, p. 4.

139. New York State Department of Health, Division of Medical Care Services and Evaluation, New York State Hospital Utilization Review Data System (undated).

140. See, for example: William Scanlon, "The 1974-75 Recession and State Financing of Health and Welfare Programs," Urban Institute Paper 986-5 (Washington, D.C.: The Urban Institute, September 1976).

141. See, in general, Clark Havinghurst, and Clark Havinghurst and Randall Boubjerg, "Professional Review Organizations and Health Maintenance Organizations: Are They Compatible," Utah Law Review, Vol. 381 (Summer 1975), pp. 381-421.

142. G. Majone, "On the Logic of Standard Setting in Health and Related Fields," in N.T.J. Bailey and M. Thompson (eds), Systems Aspects of Health Planning, North-Holland Publishing Co. (1975), pp. 279-290.

SELECTED URBAN INSTITUTE BOOKS AND PAPERS

The Urban Predicament. William Gorham and Nathan Glazer, Editors. (1976, URI 14500, 363 pp., $4.95 Paper. URI 14700, 363 pp., $10.00 Hard Cover.)

Policy Exploration Through Microanalytic Simulation by Guy Orcutt, Steven Caldwell, Richard Wertheimer, et al. (1976, URI 15300, 352 pp., $5.95 Paper. URI 15100, 352 pp., $12.50 Hard Cover.)

Time of Transition: The Growth of Families Headed by Women by Heather L. Ross and Isabel V. Sawhill. (1975, URI 12600, 224 pp., $4.95 Paper. URI 12800, $10.00 Hard Cover.)

Toward Efficiency and Effectiveness in the WIC Delivery System by Marc Bendick, Jr., et al. (1976, URI 14600, 216 pp., $4.50.)

Income Transfers and Family Structure by Isabel V. Sawhill, et al. (1975, URI 13100, 211 pp., $5.50.)

The Web of Urban Housing: Analyzing Policy with a Market Simulation Model by Frank de Leeuw and Raymond J. Struyk. (1975, URI 12700, 231 pp., $4.95 Paper. URI 12900, 231 pp., $10.00 Hard Cover.)

Para-transit: Neglected Options for Urban Mobility by Ronald F. Kirby, et al. (1975, URI 78000, 288 pp., $4.95 Paper. URI 87000, 288 pp., $10.00 Hard Cover.)

Research Priorities for Crime Reduction Efforts by Henry S. Ruth, Jr. (1977, URI 17200, 140 pp., $3.50.)

Monitoring the Impacts of Prison and Parole Services: An Initial Examination by Louis Blair, et al. (1977, URI 16900, 64 pp., $3.95.)

Growing and Declining Urban Areas: A Fiscal Comparison by Thomas Muller. (1975, URI 13400, 121 pp., $3.50.)

Fiscal Impacts of Land Development: A Critique of Methods and Review of Issues by Thomas Muller. (1975, URI 98000, 68 pp., $2.95.)

Land Development and the Natural Environment: Estimating Impacts by Dale L. Keyes. (1976, URI 13500, 128 pp., $4.95.)

Using an Impact Measurement System to Evaluate Land Development by Philip S. Schaenman. (1976, URI 15500, 96 pp., $3.95.)

The Law of the Land: Debating National Land Use Legislation, 1970-75 by Noreen Lyday. (1976, URI 15400, 72 pp., $3.00.)

Land Use and the Legislatures: The Politics of State Innovation by Nelson Rosenbaum. (1976, URI 15400, 93 pp., $3.50.)

Citizen Involvement in Land Use Governance: Issues and Methods by Nelson Rosenbaum. (1976, URI 11500, 82 pp., $3.50.)

Social Impacts of Land Development: An Initial Approach for Estimating Impacts on Neighborhood Usages and Perceptions by Kathleen Christensen. (1976, URI 15700, 144 pp., $3.95.)

No Little Plans: Fairfax County's PLUS Program for Managing Growth by Grace Dawson. (1977, URI 17100, 160 pp., $3.95.)

Program Analysis for State and Local Governments by Harry Hatry, et al. (1976, URI 13700, 155 pp., $3.95.)

For a complete list of all Books, Papers and Reprints published by The Urban Institute, please write to the Publications Office for a catalogue.

CMS LIBRARY
3 8095 00013516 6

THE URBAN INSTITUTE
2100 M Street, N.W., Washington, D.C. 20037